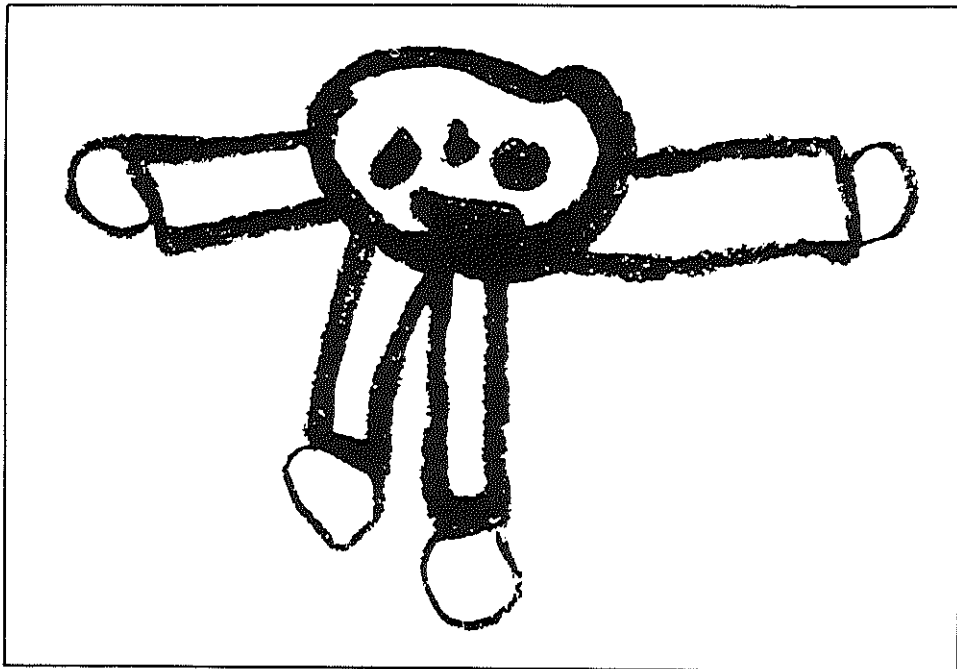


Ruth Anne Olson

**EVALUATION AS INTERACTION
IN SUPPORT OF CHANGE**



Ruth Anne Olson

**EVALUATION AS INTERACTION
IN SUPPORT OF CHANGE**

University of North Dakota
Grand Forks, N.D. 58202
December 1980

Copyright © 1980 by Ruth Anne Olson

First published in 1980

North Dakota Study Group
on Evaluation, c/o Vito Perrone,
Center for Teaching & Learning
University of North Dakota
Grand Forks, N.D. 58202

Library of Congress Catalogue
Card Number: 80-85357

Printed by University of
North Dakota Press

This series is supported by funds
from the Rockefeller Brothers Fund
and the University of North Dakota

Series Editor: Arthur Tobier

In November 1972, educators from several parts of the United States met at the University of North Dakota to discuss some common concerns about the narrow accountability ethos that had begun to dominate schools and to share what many believed to be more sensible means of both documenting and assessing children's learning. Subsequent meetings, much sharing of evaluation information, and financial and moral support from the Rockefeller Brothers Fund have all contributed to keeping together what is now called the North Dakota Study Group on Evaluation. A major goal of the Study Group, beyond support for individual participants and programs, is to provide materials for teachers, parents, school administrators and governmental decision-makers (within State Education Agencies and the U.S. Office of Education) that might encourage re-examination of a range of evaluation issues and perspectives about schools and schooling.

Towards this end, the Study Group has initiated a continuing series of monographs, of which this paper is one. Over time, the series will include material on, among other things, children's thinking, children's language, teacher support systems, inservice training, the school's relationship to the larger community. The intent is that these papers be taken not as final statements--a new ideology, but as working papers, written by people who are acting on, not just thinking about, these problems, whose implications need an active and considered response.

Vito Perrone, Dean
Center for Teaching & Learning,
University of North Dakota

Contents

	Introduction	1
1	Organizational and Methodological Frameworks	3
2	The Process of Evaluation	4
3	The Process in Operation: How It Worked in Two Programs	22
	Bibliography	28

Introduction

"Evaluation is something that we've always viewed with extreme nervousness. We always did what we were told we had to do, and then we shut the door on the whole topic as quickly as possible." Among educators, this program coordinator's view of evaluation is a common one, and with good cause. Educational evaluation typically has been a matter of outsiders looking at programs or curricula to recommend which should be continued and which eliminated. In the past 10 years, outside funding agencies have made such evaluation standard practice.

This development has been of increasing concern to educators. They question why new programs must meet rigid evaluation requirements while traditional programs continue unevaluated. They protest incidents where evaluation designs shape the programs themselves and evaluators, not program planners and consumers, determine the definition of success. They are dissatisfied with the reduction of complex interactive processes of teaching and learning to percentages and grade equivalent scores.

In the midst of these objections, however, there also runs a strong respect for information, with educators at all levels asking for more information on which to base their own daily decisions. Teachers *do* want to know what is happening in their classrooms; administrators *do* want to understand the effects of their programs. But they want the information in terms of their own questions and concerns, in ways that recognize complexities, and in forms that lead to growth instead of to hatchet jobs.

Such has been the case with a growing number of programs that have sought evaluation services that promised to contribute insights into their own processes and development. At the Prospect School in North Bennington, Vermont, ongoing evaluation and research activities have been intimately intertwined with teaching and staff development activities from the beginning of the school. A private school, Prospect has consistently funded and maintained those activities during its 14 years of existence. Marcy School, a public school in Minneapolis where I worked, developed a similar evaluation component with the help of federal funds. Unfortunately the value of such growth-oriented evaluation is made to bow before budget constraints. It is simply beyond the budgets of most schools, public or private, to assign evaluators to individual programs. What is there to do?

Ruth Anne Olson is an independent evaluator who is currently under contract with the Minneapolis Public Schools.

In 1978, the Special Education Department of the Minneapolis Public Schools decided to include an evaluator on its staff. I was appointed, and asked to design and offer an intra-department evaluation service to be made available to but not imposed on program staffs. Its purpose was to help collect information that addressed issues which program leaders, teachers and the department defined as important. The service has since become an integrated part of department activities and continues into its third year.

In many ways this task is different from my earlier work with Marcy School. Instead of concentrating on a single program, I now work with about 27. I am not on-site with any particular school and I am not directly acquainted with students, staff, issues, and parents. Instead, I move in and out of programs and always remain an outsider to the daily work of teaching and learning.

Yet in fundamental ways the task for Special Education remains similar to that developed at Marcy. As at Marcy, I am available to help staffs define the questions of concern to them, to help gather information that lends perspective in answering those questions, and to place that information in useful formats. Evaluation has become *their* tool for growth and development. The evaluator is in their service.

In this monograph I intend to describe how that happens because it is a process and a concept which others may use to influence their program planning and staff development needs. The examples and language come from special education, but the process is not unique to that context. The process is relevant for any program designed to enhance human growth and change.

In an organization where people previously "shut the door on the whole topic as quickly as possible," people now voluntarily give both thought and time to evaluation. Though not every evaluation effort has proven to be as productive as hoped, the door to evaluation is clearly open.

Organizational and Methodological Frameworks

Minneapolis, Minnesota is a city of some 370,000 residents and a public school enrollment of approximately 44,000. The Special Education Department serves about 7,000 of these students, working within the extensive state and federal laws that govern the rights of handicapped students. Services are available for students with problems related to speech, learning disabilities, emotional disturbance, social maladjustment, mental retardation, hearing and visual impairment, autism, physical impairment, multiple handicaps or situational stress due to teenage pregnancy or motherhood.

Within regular school settings, the department provides special education services for students in kindergarten through the 12th grade. These services include assessments to identify special problems, support to the regular education staff in the form of materials and activities appropriate to handicapped learners, and direct developmental and remedial instruction for students. Outside regular schools, more than 20 special-school programs provide full-time services to students aged one through 21 who require more intensive instructional support. The department also serves 14 independent hospitals and agencies and assists handicapped children in non-public schools. Within this organizational framework, the Special Education Department tries to provide a system of evaluation which yields meaningful information for decision-makers at all levels.

Accordingly the system is more formative than summative, more internal than external, more responsive than preordained. It focuses on questions of quality as well as quantity and of process as well as outcome. It recognizes subjectivity as well as objectivity and emphasizes validity more than reliability. It focuses on programs as wholes more than on components and is open to both formal and informal collection and reporting methods.

2

The Process of Evaluation

Evaluation as described here is a cooperative system. Its purpose is to be of service to program participants, who themselves play key roles in each aspect of the process. In other words, the task is not research but in-house evaluation.

Academic research is not formative evaluation. First of all, the purpose of academic research is the production of knowledge: the gathering and organizing of information to further one's understanding of cause and effect. For educators, research is a source of information about how children learn and about what organizational patterns, classroom activities, materials, and interactions will most likely enhance children's learning. Good research is designed to allow the transfer of results from one setting to another.

By contrast, in-house evaluation is grounded deeply in the here and now of a particular situation. The focus is confined to the particular students, teachers, buildings, bus schedules, community tensions, and other influences which touch on that program and combine with each other to help or hinder success. Experimental or quasi-experimental designs are out of the question, as school staffs are working with a single, defined group of students. The design for an evaluation is a limited case study. It is appropriate that the information that results may or may not have relevance to anyone outside the immediate setting.

Secondly, researchers typically spend months, even years, designing their research projects. By contrast, I am likely to get a call in September saying, "We've started this new program and would like to have some help in evaluating it." The buses are running, the students are enrolled, and the teachers are teaching. Evaluation, to be effective, needs to start immediately. In an ideal world perhaps that is not the way it should happen. But in the practical world of public education, it probably cannot happen any other way.

A third important difference between research and this in-house formative evaluation is in the staffing and resources available. In the evaluation described here, staff members are often their own data collectors. They are full-time teachers also trying to collect information. One situation provided a particularly valuable lesson in understanding the limits of such resources.

Several of the staff involved in the evaluation of a new program had studied an observation system that seemed tailor-made for their situation. We struggled with questions of timing and scheduling, and were repeatedly unable to make the system fall into place. Frustrated, a teacher and I met with the academicians who had designed the observation. We came away from that meeting finally understanding that the system was designed to use three people simultaneously, observing the 15 students six hours a week over a nine-week period. We had money to hire one observer for 30 hours in a program involving 80 students and few controls. We quit trying to adapt the professor's system to our needs, and focused instead on the practical considerations of how to get the most worthwhile information from our limited resources.

The differences between research and program-based evaluation are particularly important distinctions about which to be clear in working with teachers. Most teachers do not have experience with evaluation designed to be helpful to them, but most do have experience with research, whether because of graduate study, or having designed curriculum, or chosen a teaching method most appropriate to a situation. As a result, they tend to feel constrained by the principles of research design. They need to be continually reminded that they are not wrong to ignore those principles. They are, in fact, doing something quite different from research.

ESTABLISHING AWARENESS

The evaluator must take responsibility for initial contacts with program people. These contacts serve at least three purposes:

1. to find out people's interests in evaluation,
2. to make them aware that evaluation services are available,
3. to overcome notions they may have of evaluation as a thing apart from themselves valuable only to outsiders looking in.

The structure of the particular organization involved will guide decisions about how best to make those initial contacts. In the Minneapolis Special Education Department both individual contacts and general group presentations were desirable.

In the initial planning stages, I met with individual program coordinators to learn of their particular interests, concerns, and goals regarding evaluation. I also tried to observe programs in operation to gain a

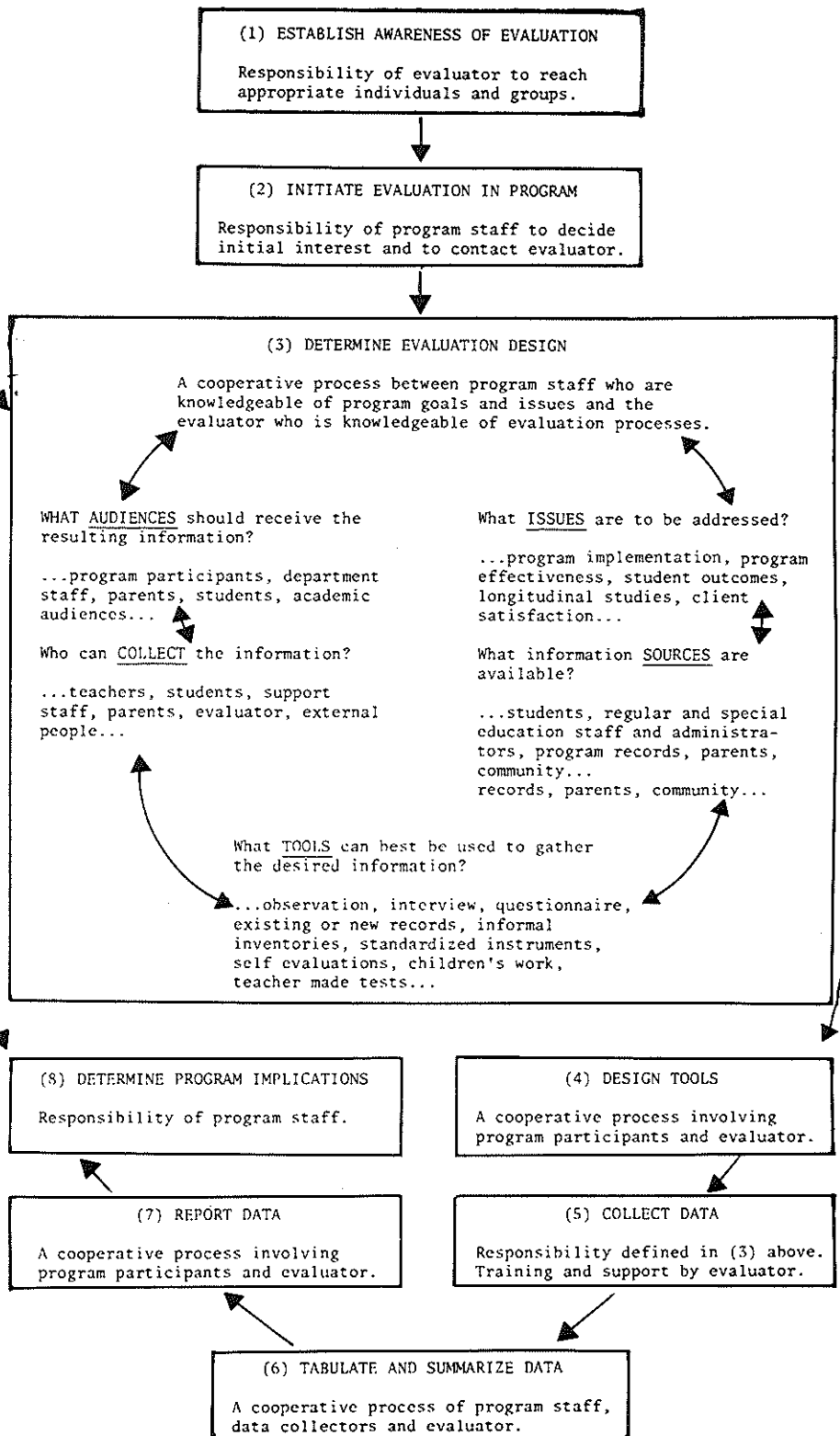


Figure 1: The Process of Evaluation.

basic sense of program functioning. Those conversations and experiences provided information about the value of previous evaluations and hints of approaches that might be helpful in the future. A few program people already had some specific tasks in mind and were anxious to get started as soon as possible.

Staff people with various levels of responsibility were selected to form an evaluation committee to serve as a sounding board for department-wide issues and planning. The committee fulfilled a valuable function at early planning stages. When the new process for evaluation was well established within the department, the committee dissolved.

In the first few weeks, I gave brief presentations about evaluation to leadership groups within the department. These early presentations focused on the proposed evaluation process, how people could gain access to it, and the kinds of questions it could help. Several months later, a second set of presentations included descriptions by staff of the evaluation processes which they had used within their own programs: how questions were defined, the sources of data, data-gathering tools, collection and analysis of data, and the uses and limitations of the resulting information.

Awareness-building is continuous as information is routinely shared about evaluation projects in different programs. Knowledge of an evaluation in one program sometimes triggers the imagination of teachers in another about what would be helpful to them. Just as important, knowledge of the limitations and failures of some evaluation designs may cause other staffs to reconsider their own directions.

INITIATING EVALUATION

Actual evaluation projects are initiated by program people coming to the evaluator with their own ideas, either specific or general. The staff roles of people who come forward with such ideas vary widely. Most often a program coordinator and teachers or other support staff decide together that they need some particular information or have a general question to explore.

For example, teachers and coordinators in two programs--one for physically handicapped senior high students, the other for emotionally handicapped junior high students--were interested to know what had happened to their graduates. They wanted to learn how their students were faring in education and life after leaving these programs and how the students thought special education had helped or hindered them in later experiences. In both these cases, the program coordinators asked my help in deciding how to gather the needed information.

In other programs, people beginning new projects wanted information that would help them assess how well the new activities were working. They wanted to know

what was happening as intended and what was happening differently, whether for good or ill. In each case, the program coordinator and staff had preliminary discussions about evaluation, and then approached me with partially formed ideas of directions to explore.

Sometimes a teacher alone wants assistance in looking at a classroom or the teacher's role within it. For example, a teacher in an assessment classroom wanted to examine the implications of her own task in assessing children. She sought ways of stepping back from classroom immediacy in order to distinguish between a child's behavior that was inhibiting learning and her own response to that behavior. This teacher asked me to help her focus and refine the half-formed questions in her own head.

In another instance, a school-based program had carried out program evaluations for several years before my coming on the scene. Their evaluation committee had defined questions, collected data, and written reports. Members of the committee approached me for new ways of collecting information or of organizing the data already gathered.

Administrators also, several steps removed from working with students, have legitimate needs for information. A high school assistant principal had been instrumental in changing the model for special education services in his building. He wanted to know the parents' opinions of opportunities available to their children. In another case, central board special education administrators had noted two teachers' enthusiasm about a program change that the teachers had implemented. The administrators needed data to assess the success of the change for the long-range, and whether it was transferable to other schools or other teachers.

In short, the call for assistance with evaluation can come from almost any staff person at any level of the organization. The only requirement is that those who ask for help be willing to accept their role in the process that follows. The importance of that cooperation should not be underestimated. The evaluation process itself heavily depends on cooperative involvement of both evaluator and staff.

DETERMINING DESIGN

Beyond a shadow of doubt, the usefulness of evaluation for any program rests on the large and complex step of determining its total design. At this point cooperation becomes crucial. The program staff have requisite information about program goals, history, issues and resources. Most important, they are the ones who will eventually use the information. The evaluator, on the other hand, is knowledgeable about evaluation processes and about other evaluation projects throughout the department.

The logistics of design are relatively simple at the start. The evaluator and a person from the program identify the people to include in the planning. In the most general terms, such a list includes representatives of all the roles that will eventually be a part of the evaluation. The inclusiveness is important. If administrative lines cross, then both administrators need to be involved in early planning. If teachers will be part of the data collection, then someone who can speak for those teachers should be present. Needless to say, touching all the bases is easier said than done. Inadvertently, but all too often, someone is omitted who should have been included--someone whose perspective, position, and ideas are important. Such mistakes can and should be quickly corrected.

As Figure 1 schematically shows, the tasks of design are highly interdependent. Each decision affects and is affected by each other in the process. For example, determining the issues depends to some degree on access to sources of information and on who is available for data collection. By the same token, identification of data-gathering tools depends on questions being asked, on the sources of information, and on the skills and time available for collecting. For sound planning, the five tasks of design must be recognized as continually interactive.

WHAT ISSUES ARE TO BE ADDRESSED?

Program staffs usually have a general sense of their particular interests. In the first year of my work with the Minneapolis Special Education Department, evaluation issues fell into seven broad categories:

1. student progress
2. implementation and effectiveness of new programs
3. colleague and client opinions
4. follow-up of program graduates
5. staff goals and growth
6. evaluation of departmental processes
7. issues cutting across the department as a whole

Student Progress. By law, every student served by special education must have an Individualized Education Plan and Periodic Reviews of Progress. To meet those requirements--as well as to meet the general expectations of responsible school systems--most programs within the department had developed systems for recording and evaluating student growth long before my presence on the staff.

Nevertheless, several programs sought suggestions for strengthening or broadening their information about students, and for more effectively organizing their routine records. In such cases, I could pass on ideas from one program to another, suggesting that a system that I knew to be effective in one place might be copied or adapted in another as well.

Sometimes I could help strengthen the validity of existing systems. One program, for example, routinely kept systematically-gathered descriptive information which lent depth and context to their understanding of students' needs. But they sought and needed my help to organize that information as qualitative statements about children's growth and learning over time.

Implementation and Effectiveness of New Programs.

New programs and changing procedures have been motivated by changes in the law, and by changing student needs. The comprehensiveness and scope of these programs vary. Several new programs in senior high schools involve 12 to 20 students per school, one or two hours per day, for one trimester. One such program integrates the services of a speech clinician, a special education resource teacher, and a regular education English teacher. A new course in four senior high schools focuses on the need to assess students to plan their high school program. A new program in the special-education preschool program groups children in four classrooms without reference to their defined handicaps. The program for hearing impaired students is finding ways to integrate all of its students in the regular elementary program for varying amounts of time throughout the school year. Other programs in the department involve changes in the grouping of children, in the curriculum, or in the responsibilities placed on staff.

In the midst of such variety and specialization, what is the evaluator's task? It is to help people in the new programs--when they ask--to be specific about what they want to know. This is surprisingly difficult. People responsible for open-ended new work often hope impractically for unlimited information. To define the particulars of any given evaluation, it is important to focus on several sets of questions.

--How can you know if the program has been effective?

Will the students have learned something they didn't know before? Will they treat each other differently? Will they solve problems differently? What will success look like?

--Whom will the program affect besides the students involved? Will staff have to assume new roles? Will another group of students lose some resources because of the focus on this new program?

--How will you benefit from answering these questions? Which pieces of information are most

likely to be useful?

The evaluator must be almost ruthless in continually urging people to be practical and to accept the inevitability that much information they would like to have cannot be gathered.

Colleague and Client Opinions. Often the views of people outside a program are important ingredients for particular evaluations within the program. Parents may be asked what they think of a new activity or a regular education teacher may be asked about the progress of handicapped students.

Beyond such auxiliary importance, colleague and client opinion may also be evaluation's prime focus. To do their own work effectively, teachers need to know what other people think of it. For example, most Minneapolis special education programs are directly involved with professional staffs outside the special education department. The perceptions of those other staffs are vital. When special education arranges tutoring for students enrolled in nonpublic schools, what do nonpublic school administrators, teachers, and parents expect from that service? What is their opinion of communication and coordination with the public system? Are they satisfied with special education's assessment and tutoring? Do they have suggestions for smoothing the transportation and organizational arrangements? Similar questions arise regarding special education services to students and teachers in the regular public schools.

Follow-up of Graduates. Natural curiosity provokes the question, "What has happened to our graduates?" By pursuing the question, teachers and administrators can look at their own programs in light of what happens to students when they leave. They can benefit from knowing areas in which students succeed or fail after leaving the program, and from knowing what graduates think of their special education experience. With such information staffs are better equipped to improve their own programs.

Data about graduates can also help identify needs for new programs. It may show that many students' needs are not being met after they leave special education programs and give reasons for exploring program additions.

Teachers in a Minneapolis special education preschool program are doing extensive follow-ups of students who have gone on to regular kindergartens. They have considerable information about the students at the time they left the preschool; now they want to know how the students have fared since, how their kindergarten teachers see the students' strengths and weaknesses, and what special education services are now available. In the same way, several secondary programs want to locate their graduates and learn about their educational experiences after leaving the special education school.

Staff Goals and Growth. Several staff people have sought help in establishing individual goals and in reflecting systematically on the roles and demands of

their professional positions. This may occur in a centralized way. For example, the coordinator of speech and language services asked that each speech clinician set self-defined goals for the school year. She suggested a broad range of possibilities and encouraged people to focus on areas which would be particularly meaningful to their individual school assignments. Each clinician was to define a goal and a system of assessing success in achieving that goal at the end of the school year.

Such reflection can also come about by individual initiatives, as a teacher seeks ways of looking at her particular role. The task of the evaluator, in that case, is to suggest ways of clarifying the questions tumbling through the teacher's head and to suggest ways the teacher can step back from immediate experience and reflect systematically on what happens in the classroom. Typically, other teachers are soon drawn into the same questions. The process may serve as a focus for peer-group reflection and growth.

Departmental Processes. Though process questions are often part of broader evaluation issues, they can also stand alone. One example is a study of referral procedures for secondary school students leaving special schools to return to their regular high schools. Questions center on communication between the special and regular schools, program options for students with special needs in the regular school, and the responsibility for follow-up to assure the most successful transition possible. With some evidence that not all students are making a successful transition, a department administrator wants information which can be used to strengthen the process.

Department-Wide Issues. The last category of evaluation issues includes questions which cut across many or all the programs within the department. One such issue concerns the concept of Least-Restrictive Environment (LRE). By law, a handicapped student must be placed in the program which least restricts without detrimental effect. The central leadership of the department has chosen to focus a department-wide evaluation on implementation of LRE policies. Questions center on defining "least restricted," identifying the levels of restrictiveness within the programs, determining the relevant factors for placement of students, and following the movement of students from one level to another.

In no way does this list of topics exhaust the evaluation questions which might legitimately be pursued. The basic question for any program is simply, "What is it you want to know?," and then, equally important, "What will you do with the information once you have it?" The latter question is the acid test. There is no end to things people would like to know. In this process of evaluation, however, there is little time to collect information just because it is interesting. That is a

waste of people's time. Once they have wasted their time that way, they will be hard-pressed to see the value in tackling any future evaluation questions.

WHAT SOURCES OF INFORMATION CAN BE TAPPED?

To some degree, the answer to this question may be inherent in the issues addressed in the evaluation. However, one must not assume too much too quickly. Few issues in education have single answers. For that reason it is important to broaden, rather than narrow, the sources from which information is gathered.

Most parents and educators recognize, for example, that there is no easy and fast way to determine what a child has learned. We may know what we have shown or demonstrated to a child, but what that individual takes away from the experience, how she has absorbed it, where she will try it out, and what pieces of it she will later remember are all questions very difficult to answer. By the same token, who is to say whether a new program is successful? The parent of a child who is doing well, the parent of a child who is distressed by the change, the teacher whose role is now different in major ways, and the teacher whose responsibilities have increased because of assignments of some of his colleagues to the new program--all may have very different answers to the question of success. For these reasons, it is essential to consider the many people who have legitimate views of the issues being considered.

In the planning process, active consideration should be given to the following sources:

1. students
2. parents
3. professional staff, including special education and regular education staffs, teachers, social workers, teacher aides, psychologists, administrators, tutors, and others.
4. community people, including volunteers with the program, employers, recreational supervisors, juvenile court staff, and others.

The goal is to construct a collage, a rich construction of opinions, reactions, achievements, and experiences from which to draw conclusions about the interplay of people and accomplishments.

For example, the staff of an elementary program for multiply and profoundly handicapped children began some activities for integrating their students in swimming, lunch, and activity programs in several regular schools. To gather information about the affects of that program, the staff chose to gather information from parents, from

regular education teachers, social workers and administrators, from special education staff involved in the program, and from students.

It is not practical, of course, to gather information from all possible participants. Judgments must be made as to whom to include and whom to leave out. Who are the most essential participants? Is the number of students (or parents, or teachers) large enough to justify including only a sample of the total or must all be included? And--probably the most important question of all--what are the data-collection tools and how time-consuming will it be to gather information from any given source? At this point, of course, many different factors begin to tumble on top of one another. All the pieces related to sources, tools, data collectors, and audiences must be played with like a jigsaw puzzle, to determine the ways that things fit together, and finally to arrive at the design which feels most comfortable and which makes the most sense in terms of the reasons for doing the evaluation.

WHAT EVALUATION TOOLS CAN BE USED?

As with data sources, the broadest possible range of data-collection tools should be considered. The list of possibilities includes:

1. Observations, formal and informal
2. Interviews, individual and group, structured and reflective
3. Questionnaires
4. Program or student records
5. Informal inventories of skill growth
6. Standardized instruments
7. Self-evaluations
8. Children's work
9. Teacher-made tests
10. Journals

Evaluation planners should be made aware of the potential of each of these tools. For example, if teachers choose to examine student progress, they might do so by observing students' use of materials, involvement in activities, or interactions with one another or with

adults. They might interview students or parents about their perceptions of growth and progress, administer informal inventories to test reading ability, ask students to complete self-evaluation forms, or assess the skills demonstrated in children's work.

Recordkeeping procedures should be organized to make data collection a continuous process for providing planning information. Certain records may already be kept and need only to be reorganized or summarized in a way to make them more useful.

Major considerations in selecting a tool should include whether it is feasible to use the tool, and whether the tool fits the dimensions that program people have identified as important. Testing and measurement experts have standards by which data collection tools are judged to be reliable (i.e. whether a given tool provides similar results even when used by different people), and are judged to be valid (i.e. whether the instrument measures what it claims to measure). Standards of validity are relatively low for some kinds of instruments--a situation which I find quite puzzling. If I want to know the temperature on a Minnesota day in order to decide whether or not to wear long-johns, I will not choose a thermometer which the manufacturer says will give me readings 70 percent determined by temperature and 30 percent by uncertain quantities of air pressure, radioactivity, ultraviolet rays, or undefined factors. By the same token, I don't believe educators should make judgments about programs and students on the basis of instruments only 70, 60, or even 50 percent valid for the skills their labels say they measure.

Most systems of data collection are less than perfect. It is important, therefore, not only to seek out a variety of sources, but also to use a variety of tools --to balance numerical and descriptive information in a way that allows a triangulation of information.

WHO WILL COLLECT THE DATA?

Students can serve as data collectors, particularly if one of the tools is classroom records of activities and growth. Having children chart their own growth not only saves the teacher's time but serves as an effective learning tool and motivator for the students.

Several schools have parent-staff evaluation committees in which participants design their program evaluations and share the tasks of collection. Teachers are the natural people to collect some kinds of information throughout the school year; parents can step back and observe certain classroom activities free of the responsibilities that the teacher has. Teacher aides, counselors, social workers, and other support staff are the logical people to put together other kinds of information.

Because of the particular issues being examined, the sources of information, or the tools of collection,

it is sometimes important to have someone completely outside the program serve as data collector. That may or may not cost extra money. I can often be responsible for some small pieces of a larger evaluation plan. If three or four people need to be interviewed or if a few hours of observation information is to be collected, it may make most sense for me to do that.

In some cases, a staff wants to collect extensive information which demands a major commitment of time. We may arrange for volunteer help from a student at a nearby college or university. Such students can get the job done and in turn receive experience in evaluation or data they need for class assignments. In other situations, the only solution to the data-collection needs is to find money in the budget and to hire someone.

Regardless of who collects the information--but especially if it is someone from outside the program, everyone involved needs to be clear about what pieces of the information can be made public. Collection of valid data requires trust between the collector and the participants. Participants must be confident that information will not be used inappropriately.

WHAT AUDIENCES SHOULD RECEIVE THE INFORMATION?

The essential question here is, "Whom do you explicitly want to know what you've found out?" Since the information will be used for program planning, program participants must receive it. Beyond that, is there reason to report it to central administrators, to parents, to students? Would it be desirable to share the information with academic audiences or professional organizations? What parts of the information should be shared? Will we describe only the evaluation process or also report the specific findings? Will we report an overall summary or will we tell others the specific details?

These questions have no automatic answers. For us, the governing factor is the purpose of facilitating program planning and growth. We release evaluation information only insofar as it serves that purpose. This whole process, remember, is *internal* evaluation; *external* reporting is a secondary concern.

When we have juggled the pieces and decided how to fit them together, we have an evaluation design. In many ways, the hardest part is over. The work of carrying out the design is still to be done, but everyone at least has agreed on what the job is.

DESIGNING TOOLS

There is something exact and satisfying about reaching this stage of the evaluation process. There are still hard decisions to make, but the reward of concrete accomplishment is in sight.

The first step in designing the tools is to explore the use of systems that others have already developed. Within the program itself or within other programs in the school district there are likely to be surveys, observations, or recordkeeping systems which can be relevant to the task at hand. Various state and national information networks are further sources of designs and tools for evaluation. One must remember, however, that the purpose of these evaluations is to gather information particular to certain settings. It is unlikely that tools developed elsewhere can be used as is. They may yield ideas and directions, but almost certainly will have to be modified.

Basically, it falls to the program staff who have initiated the evaluation to identify the information they want, from whatever sources, and then set priorities. If the evaluation design includes interviews with parents or with administrators, I meet with a few of them to find out the questions they want to have included.

After everyone's ideas are collected, I draft the format of a questionnaire, an observation schedule, a recordkeeping form, or whatever tool is required. That draft is given to the people who contributed their ideas. Have I correctly understood what they were suggesting? Have I captured the essence of what they think is important? What about the format? Does it seem reasonable, useable? And most important, if they use this format and if we gather all this information, will it tell the program staff anything that is important to know?

This can be a tough time for the evaluator. One must hold one's ego firmly in hand and be willing to take some hard criticisms. Almost without fail, I have found that I misunderstood an issue that someone had raised, or in subtle ways managed to misrepresent the emphasis people wanted. Probably the most helpful experience I have had at this stage of tool development was with several parents who were reviewing the draft of a questionnaire. Confident that I had eliminated all educational jargon from my questions on the draft, I showed them my list. They picked it apart question by question. "Why ask," they wondered, "'Do you feel comfortable contacting the school when necessary?' when you can simply say, 'Do you feel okay about calling the school when you need to?' 'Do you receive adequate information about your child's progress in school?', can be stated simply, 'Do you get enough information about how your child is doing?'

"The format, too, is all wrong," they told me. I had drafted a format with a scale of 1 to 5, labelled "very satisfied" to "very dissatisfied." They told me that such a numerical scale would be unnecessarily confusing and intimidating to some people. They suggested, instead, simple yes-no-uncertain or always-sometimes-never answers.

I revised the questionnaire incorporating their suggestions. We mailed it to 161 parents of mentally

retarded children: 61 percent returned the anonymous questionnaire without a single follow-up reminder. That rate of return is high for the particular population and for the kind of issue involved. I am confident that the major factor in the willingness of those parents to complete the form was the fact that it contained questions and directions that were easily understood.

Whatever the tool, it is essential that program people be confident it provides the information they want. That may require many drafts, many revisions, and much piloting. All that must be done, for without a good tool, everything else is for naught.

COLLECTING DATA

At this point, it is well to pause and make sure that everyone who should be informed about the data collection is, in fact, informed. Do all the relevant people know what information is being collected, by whom, and for what purpose?

In some cases, data collectors may need training. They need to be comfortably familiar with the setting in which they will work, with the program, with the classroom, with the larger evaluation design. They need to be thoroughly grounded in the tools that they will use--with appropriate techniques and cautions about interviewing, observing, or whatever. They should have participated in several trial runs and a discussion of how those went. They should know the rules about confidentiality and public information.

Careful scheduling is another necessity. We try to keep participants informed about what data will be collected from whom, by whom, and when. Inevitably, the schedule will change since the unexpected always happens; when it does, the important people are notified. If several people are involved in the data collection, some one person needs to be coordinating it all.

One aspect of data collection deserves particular comment. People are surveyed to the point of annoyance. I knew that with certainty several years ago when I was in a Rocky Mountain wilderness area and a person carrying a clipboard appeared on the trail ready to ask about my past, present, and intended future use of the wilderness.

Of course, we each think that our own survey is important. This is one that people certainly *should* answer. The key to making that happen is to make the survey important to the people whose answers we want. A step in that direction is to give respondents information about the results. I often make this kind of statement in a cover letter: "If all questionnaires are returned by January 15, you will receive a summary of the results by February 15." Then I must keep that promise.

TABULATING AND SUMMARIZING DATA

As described earlier in the comments on evaluation design, the intent of the process is to form a collage of information from a variety of sources with a variety of perspectives. As with a visual arts collage, the information collage has to be carefully constructed in order to have meaning to the viewer. Qualitative and quantitative data must be pulled together in such a way as to enrich each part, and to enhance the ability of the program staff to see the meaning and impact of each perspective represented in the total design.

I try to avoid dividing the tabulation and summary tasks into too many pieces or among too many people. The process of summarizing--for both quantitative and qualitative information--is largely one of seeing the patterns emerge; of being sufficiently involved with all the information to be able to see the sense of the whole. This is true particularly with qualitative information. For that reason, it is important to have at least one person directly in contact with the numbers and comments from each area of data. Other people can and should be a part of this process. For example, all data collectors report subjective impressions gathered in the course of their work; several people should ponder the suggested patterns and mentally mull over the relationships of some pieces to others. But the resulting picture will have greater depth if at least one person is in touch with it all.

At least two cautions are in order. The first is about statistical analysis. The purpose of statistical tests is to clarify the meaning of the numbers by relating them to accepted standards of significance, relationship, and other dimensions. For people who are not familiar with the language of statistics, however, such tests can cloud the issues rather than clarify them. For that reason, I strictly limit statistics to those calculations which program people will easily understand.

The second caution relates to maintaining an appropriate perspective on the profundity of the information collected. When things get down on paper, they seem to assume an importance and an aura of truth which they may not merit. Numbers can have an especially insidious effect. Statistical analysis of quantitative information and the quantification of qualitative information can often lead to an inappropriate simplification of complex perspectives. Everyone involved must recognize those limitations.

REPORTING INFORMATION

Reporting can take several forms and, as with every other decision related to this process, the choice of format depends on what is of most direct use to the program.

People who were asked to respond to questionnaires or who agreed to be interviewed should receive a summary of the results. That summary generally need not be longer than one or two pages and contains such information as the design of the study, the percentage of respondents, a gross summary of the resulting information, and an invitation for recipients to request a more detailed reporting of results if that is appropriate.

Program participants who actually intend to use the information require a more detailed review. It is usually most successful to give written information first, then to meet to discuss the context and limitations of the study and its emerging patterns, meanings, and questions. I try to be prepared to point out some patterns in the information to help people begin looking at it. But in no way do I--nor can I--presume to be definitive about the importance of the information gathered. It is program participants who will be most insightful about what they see there. It is they who live the program to the degree where they can best know the information.

Whatever the form of reporting, I give public credit to everyone who assisted in the design, data collection, and reporting. Those people found the project to be of enough interest and importance to give their time and ideas to it. Listing their names on any resulting report or summary is minimal recognition of that involvement.

Program staff often wonder whether they should write a complete report of the study and its results. The answer rests on the recognition that the purpose of the study is program planning and development. If a large report will further that purpose, then it is worth writing.

DETERMINING PROGRAM IMPLICATIONS

With all the information before them, with patterns and trends identified, the program staff who initially requested the evaluation can return to the issues which they originally identified and ask themselves, "Does all this suggest that we change anything?" From here on the responsibility clearly falls on program people. At this point I largely step out of the picture unless invited by the staff to be part of their discussions.

In this department there have been many visible program changes as a result of information collected in program-requested evaluations. In several programs, procedural changes were made to correct problems that had been identified. Decisions have been made to expand programs which were seen as successful. In one situation, a program had not been successful in accomplishing all that was intended, so expansion was accompanied by procedural changes to address the areas of weakness.

In several situations, the evaluation revealed areas in which staff development activities could

contribute to growth and positive change in the program. Some problems could be addressed best by opening new channels of communication.

Two evaluation projects involved large amounts of time and a large number of people but have not, to my knowledge, been used in a single program or staffing decision. There is something to be learned from those; they reinforce the basic premise of what makes this kind of evaluation meaningful. In both cases, the motivation for the evaluation came from outside the immediate program staff and the program staff saw it as something apart from themselves--something they were doing for someone else. The particular circumstances of the two differed, but they shared that sense of externality. Nevertheless, it may have been appropriate to do these evaluations. It was important to have the information for the external audience. But the experience confirms the premise that evaluation done primarily because of external demands is likely not to have internal impact.

Besides influencing programmatic and staffing decisions, the type of evaluation described here has another important, albeit more subtle effect. The very process of requesting, designing, and carrying out a program evaluation demands that people sort out and be clear about issues which might otherwise stay muddled or slip by unnoticed. I compare this benefit to the demand for clarity which I have felt as I write this monograph. The need to be explicit, to be able to explain and defend a process, results in subtle but very real insights into that process. One must think through how one part fits with another. One is forced to ponder the things which didn't work and what was different about those which did.

That act of clarification--of reflection--is, I would argue, the real strength of this evaluation process. The demand to define issues, to identify sources of information, to design ways of documenting change and growth, all keep a program staff aware of themselves, of their purposes, and of how their work all fits together in support of children's learning. That process brings subtle changes which are as profound in their impact as any of the explicit procedural corrections that are made.

The Process in Operation: How It Worked in Two Programs

In some measure, the foregoing description of the many decisions and contacts that are made and the steps that are taken to carry out a useful evaluation portrays an ideal: the way things *should* happen, and the benefits that *can* result. In actuality, many big and little things get in the way of that ideal.

Some hindrances may be organizational. Evaluations often cut across special and regular education, preschool through post-high school education, public and nonpublic, or other administrative divisions, obscuring the appropriate channels of communication. Everyone who should know about the evaluation must be kept informed, but at the same time we want to minimize our contribution to the mass of unnecessary communications that flood people in large institutions. It is sometimes hard to know what is necessary and what is extraneous.

Another problem is that of priorities. Information that seems important to one individual or group, may be of low priority to others who need to be involved. In addition, priorities change. An evaluation may seem important to the participants at the start, but relatively much less so by the time the data are collected.

Most often, the problems have to do with ordinary human error. With evaluations of several programs taking place at any given time, either I will manage to forget some detail, or someone else will fail to finish some task that must be done before the next meeting. Many such problems can be avoided by being organized and careful. I doubt very much that they can be eliminated altogether.

To give the reader a sense of the flow, of how the ideal translates into the real, I will describe three evaluations in detail. The three differ in several dimensions. Together they illustrate a variety of strengths and vulnerabilities.

PRESCHOOL FOLLOWUP

Establishing Awareness and Initiating the Evaluation

The special education preschool program serves approximately 130 four- and five-year-olds who have been classified as emotionally handicapped, developmentally delayed,

speech- or language-delayed, and hearing- or visually-impaired. The program staff first learned of the availability of evaluation services when I interviewed the program coordinator about her staff's evaluation interests.

Several months later the coordinator asked for help in constructing a study to follow up on children who had left their program, and explained the background circumstances. The previous spring, the preschool teachers had identified approximately 20 children who they believed would not be successful in a regular kindergarten classroom, but who in their judgment were not ready for placement in traditional self-contained special education classes.

The preschool staff met with administrators to consider establishing a transitional kindergarten for these 20 children. Ultimately, the administrators decided they did not have sufficient evidence of need to justify the establishment of a special kindergarten, and the children were placed in regular classes.

Now, the preschool staff wanted to know if that had been a good decision--whether the question of a transitional kindergarten could comfortably be closed or whether it should be reopened. Several preschool staff people volunteered to construct a study to collect relevant information.

Design and Data Collection. Many, many meetings during lunch and after school followed in the next four months. The design issues seemed so complex that there were times when we came close to abandoning the task altogether.

Early in the process, the staff established three purposes for the study. They wanted information which would help them know whether they should

1. reopen the question of establishing a transitional kindergarten for children with certain needs,
2. consider changes within their own programs to prepare the children more adequately for kindergarten, and/or
3. explore the advantages of selective placement of some of their children, i.e., of placing children with particular needs in particular types of kindergarten environments rather than in the kindergartens of their neighborhood schools.

We considered students, parents, and kindergarten teachers as sources of information and explored various data-collection tools. Immersed in indecision, about the only decision we were able to make quickly was to eliminate questionnaires as a potential tool. Questionnaires would not establish rapport and would be too

constrained to get at the depth of information which we sought. We eventually decided to:

- Interview each child's kindergarten teacher to determine the teacher's perspective of the progress of each child.
- Observe each child to provide an outsider's perspective of the child's functioning.
- Observe each kindergarten classroom to obtain a profile of the classroom environment.
- Review available standardized test scores.

The decision not to gather information from parents was uncomfortable for everyone on the committee. Faced with limited resources, however, we believed the decision to be necessary. Some of the preschool staff would serve as data collectors; they could make time for the task. Observations could be assigned to people who did not know the child in question. That was important for the sense of objectivity.

We had repeated difficulty deciding whom to include at the various stages of planning. Preschool staff and their administrative supervisors appropriately were included from the start. Then, well into the planning, we realized we should also include people who worked with special education services in regular schools. When asked to participate, they preferred to be kept informed rather than actively involved. We never did invite regular education people to participate in the planning. In retrospect we should have; it seems so obvious now, but we did not think of it at the time.

We searched the evaluation literature and tools developed for similar projects in our school district, yet we could not find a child-observation instrument that was directly relevant to our study. By now it was early spring and we simply did not have time to design our own observation tool. So we dropped our intent to observe the children and focused instead on ways of interviewing teachers and observing the classrooms.

Our teacher interview drew heavily on a questionnaire which had been developed a year earlier by the Minneapolis Department of Research and Evaluation. We asked questions about the child's affective, social, and academic functioning in the classroom; about special education services received by the child during the year in kindergarten; about the teacher's opinion of the child's readiness for promotion to the first grade; and about standardized test results available for the child.

Our way of observing the classroom environment was based on a study of open schools in Massachusetts. The observation contains four subscales which we believed would effectively differentiate among the wide range of kindergarten classrooms.

Plans for preschool staff to serve as data collectors eventually had to be dropped; it was later in the school year and the teachers' time was no longer as flexible as it had been in the fall. The study planners also believed that the information would have greater credibility if it had been collected by someone other than the preschool staff.

We explored the use of university students as data collectors and the use of special-education lead teachers* who provide technical assistance to regular schools. Both ideas had to be abandoned because of logistical problems. In the end the administrator of special school programs provided money to hire a data collector from outside the schools. The person hired already had extensive experience in schools and with interviewing; she needed no training except in regard to the specifics of the study.

As we were getting ready to begin data collection, a department administrator suggested a major expansion of the study. She saw value in having information not only on the children identified to be at-risk but on an equal number of children from the same preschool program who had not been identified as being at-risk. We received additional funding and doubled the number of children in the study.

The study now included 35 children enrolled in 28 kindergarten classrooms throughout the city. We contacted the principals and teachers and found them all not only willing to participate but in most cases even enthusiastic about the study. Finally, after months of planning, the interviews and observations were scheduled and completed by the independent data collector. There were few problems.

Reporting the Data. When the data were in and partially tabulated, several of us met to review the results. Late in the process, in fact after most everything was done, I realized that we had ignored an important source of information--namely the progress reports written at the time these children left the preschool program. I studied those records to compare teachers' reviews of each child when leaving the preschool with the information collected from the interviews of their kindergarten teachers a year later.

The pre-school staff received a written summary and a verbal presentation reviewing the study. Their subsequent discussions led to some explicit changes in the pre-school program to better prepare children for kindergarten. It also led to a city-wide meeting of kindergarten teachers and pre-school teachers. Such a meeting had never before been held in the district and was the beginning of further sharing between those two groups.

Special education administrators received a report which highlighted data bearing on the question of a transition kindergarten. The administrators did not believe that the information was decisive enough to merit such a new program, but encouraged the pre-school

*Lead teachers are experienced teachers who are on special assignments as consultants to classroom teachers.

to continue collecting data in the next school year. That was done, and the findings have precipitated active planning in increased programming for selective kindergartens.

People outside the district were also interested in the preschool study. A university committee asked if they could see the results because of their relevance to issues which they were studying. Preschool people were confident that release of the information to this committee would have no adverse affect on their own use of it. So after going through appropriate school district channels we sent the data summary to those who had requested it.

In all, it took eight months for planning and executing the study. As described earlier, it addressed three questions. The first question--whether to establish a transitional kindergarten--resulted in no further planning except to repeat the study a second year. Though we didn't know it then, two years of data would lead to action on this question. The second question--whether changes were needed in the pre-school curriculum--was answered affirmatively and led, further, to new communication between pre-school and kindergarten staffs. For the third question--possible advantages of selective placement--the data revealed no clear patterns of greater or lesser success of children according to styles of classroom environment.

On the negative side, we had problems with the selection of data-collection tools. It was difficult to find appropriate instruments and there was insufficient time to create new ones. As stated earlier, one dimension of the study had to be dropped for lack of an appropriate instrument.

Another problem lay in the number of administrative lines which this task crossed. One purpose of the study was to provide information for a reconsideration of the need for a transitional kindergarten. We stumbled several times in deciding whom to include in the planning and never did think to include kindergarten teachers and principals. We would correct that error for the second year of the study, thus making sure that planning perspectives of people other than special educators were included.

A third weakness of the study was that we neglected to establish, with the various administrators involved, clear criteria for judging whether or not the data was significant enough to merit planning for new programs. The lack of such criteria resulted in one group of people arguing that the data revealed a clear need for change and another group arguing that the data was not strong enough for such change. The need for criteria would be addressed in planning for the second year of the study.

The most important strength of the study came from the conviction held by many people that the study was

worth doing. Though the precise questions on which to focus and the ways to collect the information were often hazy and confused, the central issue was always seen as valid and educationally important. On top of their regular teaching responsibilities, the pre-school staff voluntarily assumed major responsibility for planning and designing this study. They believed it was a task that would lead to opportunities for their students and strengthen their own abilities as teachers. Kindergarten teachers and principals gave willingly of their time. Though not involved in the planning, they could see that the study would be meaningful to the children they taught and to the programs which they administered. In short, the study had direct meaning and significance to the people involved. It was not external or academic; it came out of their own needs and spoke to their own interests.

SECONDARY REASSESSMENT CLASS

A very different evaluation process was used with a special class for high school students. The Minneapolis Special Education Department provides services to approximately 800 students enrolled in regular secondary schools. Services include assessment of student needs, provision of special materials and activities to regular classroom teachers, and direct educational and vocational services to students.

The coordinator and staff of these programs had identified a problem in planning programs for students who had received special education service in their elementary or junior high schools and who were newly entering high school. Class schedules for these students typically were planned on the basis of information available from their elementary schools. There had been little opportunity to do the reassessment which would describe the students' current needs more completely and result in a well-planned high school special education program.

Special education staffs of four high schools volunteered to design and teach a reassessment class which would include orientation of the students to the school, plus observations, testing, and discussions with students so that appropriate plans could be developed for high school. Handicapped students new to the school were scheduled to take the class one hour each day for one grading period. The class, for which students would receive graduation credit, would result in a high school Individualized Educational Plan (IEP) and a tentative class program for the student's high school years. The class would also serve to establish a relationship between the student and the student's staff advisor and to establish communication between the staff and the student's family.

The administrator in charge of secondary school-based special education services had become aware of

evaluation services when he served on the department-wide evaluation committee and had requested evaluation help with several earlier projects. He now wanted to evaluate the reassessment classes and asked me to help design and complete the evaluation.

Design and Data Collection. We decided to focus the evaluation on whether the purposes of the classes were accomplished and on identification of unexpected problems and benefits of the classes. In talking with the class teachers to get their ideas for the evaluation, we found that they were most interested in finding out students' and parents' opinions of the class; we knew that we also wanted to know teachers' opinions of the classes. After two planning meetings, we decided on the following data collection:

1. Interviews with parents of students enrolled in the courses at all four schools;
2. Optional student interviews;
3. Examination of student files to determine completion of assessment, Individualized Educational Plans, and tentative high school plans; and
4. Group interviews with teachers at the beginning and end of the grading period.

Several people took responsibility for the data collection. Teachers at one of the schools designed the parent and student interviews and made them available to the other schools so that all four could use a similar format. We did not include any parents or students in the design of those tools as their intent was too cursory to merit that. The support staff person and I identified criteria for judging the assessment and student plan information in the files; she took responsibility for the actual file search. She and I planned the group interviews with teachers and shared the responsibility for conducting those meetings. I was responsible for summarizing the results of the meetings.

Our aim was to complete the data collection within two weeks of the end of the class. It became a bit complicated because so many people were involved, but we did finish close to our target date.

Reporting the Data. Most of the information was shared orally in a staff meeting. The only written report which we believed necessary was a summary of the two group interviews with teachers. That written summary was sent to all the participating staff with the invitation for them to suggest corrections or modifications.

Overall, the collected information provided evidence, from many different perspectives, that the course was worth continuing and worth expanding to other schools where the staffs were interested. The information did identify several areas for further planning and staff

development, notably a need to integrate various types of student assessment information and to communicate it more effectively to students and parents. The staff has decided to focus on these two areas as they develop plans for next year.

Relative to the previously described followup of preschool students, this study was much more limited in almost every dimension. It involved fewer teachers and touched the turf of only one administrator. The time frame was defined by the length of the course being offered, and no additional funding was needed for data-collection.

On the negative side, there were two clear problems in this study. One was the involvement of so many people in the data collection. The teachers were responsible for interviewing parents and students, but had other obligations that took precedence. The interviews did not yield particularly interesting information so there was not much motivation for teachers to make the time to do them.

The second problem was in the difficulty of summarizing the two group interviews with staff. In those interviews, staff enthusiastically discussed some results of the course which were not directly a part of its original purposes. But in summarizing those interviews, I was not always successful in presenting a fair balance of that enthusiasm with the problems the teachers reported in accomplishing the defined purposes. Fortunately, I sent drafts of the summary to everyone involved and asked for their comments and corrections. Giving people a chance to react and make suggestions was a particularly important step in this case.

The general strength of this study was that the issue on which it focused was relatively precise and contained. The program staff had defined their purposes and their question was whether those were met. The potential uses of the information were clear to the participants. They would use it to help plan for next year to determine areas of administrative support, of staff development, and of staff planning time. Overall, everyone knew what was expected in the evaluation study and they could be relatively precise in accomplishing those things.

TEACHER REFLECTION

Evaluation services are available to individual staff members as well as to program groups. One project started when a teacher asked for help in reflecting on her role in determining whether individual pre-school children are handicapped and would benefit from special education services. She and I met several times over a period of months. At first, the task seemed impossibly unclear and nebulous. She was dissatisfied and uncertain

about something in her work with children. She had a sense of what that was, but found it very hard to verbalize that sense. Our discussions gave her both motivation and form for becoming clearer about her own ponderings. I gave her informal formats for observing the activities in her classroom and reflecting on her own role in relation to those activities. She shared her reflections with colleagues and was excited to learn that her thoughts made sense to other people and touched concerns which they also had.

Some firm structure began to emerge from our discussions--a structure that made exciting sense to this teacher and might well relate to other people, too. We invited other teachers, a psychologist, a speech clinician, and a social worker to join our discussions, and they enthusiastically tested out the one person's reflections against their own roles and perceptions of children.

The process continues to snowball. Two teachers from another school were invited to join the group, and they all applied for staff development money to set their work on paper. At this point, they have developed a format for observing children and for judging special needs that seems to have meaning in several professional roles and across pre-school and elementary-aged children. They are writing about their experiences to share with other staff in the Special Education Department and hope to include teachers from other settings as they continue to explore the concepts.

This experience illustrates the power of evaluation from the standpoint of the individual. The process started with one person who wanted to examine her role and evaluate her own procedures. It has expanded to become an ongoing discussion of people, in various staff roles, who want to examine, reflect on, and grow in their work.

BEING GOOD ENOUGH

These three examples are intended to translate the flow of evaluation shown in Figure 1 into the real terms of people's interests, time constraints, indecision, and program pressures. Many aspects of these three projects were at odds with the ideal of orderly rational progression. Nevertheless the evaluations were effective in that they provided important information to the people who sought them in the first place. Figure 1 may be idealized, but it is clear that it represents a basic flow that is good enough and sound enough to sustain a great deal of damage without being rendered ineffective.

Each study required time commitments by the program staff; no one could expect that someone else would do the work. There, in fact, lies the strength of the process. Teachers define the questions, figure out how and where to find the answers, and then set about doing it.

They are not eager to waste their time, so they make very sure extra tasks are worth the effort.

Furthermore, the range of people's interests seems to have no bounds. As people try projects and experience the success of growth and change, they become more sophisticated and complex in their explorations. It is this process that yields learning and in turn encourages us to try to support the children's learning at deeper and deeper levels.

Traditionally we see teachers as being the people who teach and evaluators as the people who watch the teaching. Learning does not acknowledge such dichotomies. Schools should make evaluation a shared learning experience.

Bibliography

- Aldrich (now Olson), Ruth Anne, "Innovative Evaluation of Education," *Theory into Practice*, Vol. XIII, February 1974, pp. 1-4.
- Bogdan, R. and Taylor, S. J., *Introduction to Qualitative Research Methods*. New York: Wiley, 1975.
- Bussis, Anne; Edward Chittenden and Marianne Amarel, "Alternative Paradigms in Educational Evaluation," Educational Testing Service, an unpublished paper, October 1974.
- Carini, Patricia, *Observation and Description: An Alternative Methodology for the Investigation of Human Phenomena*, North Dakota Study Group on Evaluation, 1974.
- Cohen, Monroe (Ed.), *Testing and Evaluation: New Views*. Washington, D.C.: Association for Childhood Education International, 1975.
- Eisner, Elliot, "The Perceptive Eye: Toward the Reformation of Educational Evaluation." Invited Address at the American Educational Research Association, March 1975.
- Engel, Brenda, *Informal Evaluation*, North Dakota Study Group on Evaluation, 1974.
- Frymier, Jack, "Who is Accountable to Whom for What?" Invited Address to Association for Supervision and Curriculum Development, 1974.
- Mishler, Elliot, "Meaning in Context: Is There Any Other Kind?" *Harvard Educational Review*, Vol. 49, No. 1, February 1979, pp. 1-19.
- Olson, Ruth Anne: see Aldrich.
- Parlett, Malcolm and David Hamilton, "Evaluation as Illumination." Occasional Paper from Centre for Research in the Educational Sciences, University of Edinburgh, 1972.

- Patton, Michael Quinn, *Utilization-Focused Evaluation*.
Beverly Hills and London: Sage Publications, 1978.
- Perrone, Vito, "Alternatives to Standardized Testing,"
The National Elementary School Principal, Vol. 54,
July/August 1975, pp. 96-101.
- St. Paul Open School, "Year Six Evaluation, 1976-1977,"
St. Paul, Minnesota.
- Stake, Robert, "Responsive Evaluation of Educational Programs," unpublished statement, 1975.
- Zimiles, Herbert, "A Radical and Regressive Solution to
the Problem of Evaluation." A Paper presented at
the Minnesota Round Table in Early Childhood Educa-
tion, Spring Hill Conference Center, Wayzata, Minne-
sota, 1973.

The monograph series of the North Dakota Study Group on Evaluation

African Primary Science Program: An Evaluation *by Eleanor Duckworth, \$5*

Children's Language and Thinking: A Report of Work-in-Progress *by Edith H. E. Churchill and Joseph N. Petner, Jr.*

A View of Power: Four Essays on The National Assessment of Educational Progress
by Paul Olson

Developing Hypotheses About Classrooms from Teachers' Practical Constructs
by John Elliott

Children's Interactions in Traditional and Nontraditional Classrooms
by Sylvia Ross, Herbert Zimiles and David Gerstein

Psychological Effects of Open Classroom Teaching on Primary School Children:
A Review of the Research *by Robert A. Horwitz*

Special Education: The Meeting of Differences *by Steven D. Harlow*

A Handbook on Documentation *by Brenda Engel*

Informal Evaluation *by Brenda Engel*

An Open Education Perspective on Evaluation *by George E. Hein*

Alternative Evaluation Research Paradigm *by Michael Quinn Patton*

Observation and Description: An Alternative Methodology for the Investigation
of Human Phenomena *by Patricia F. Carini*

The Word and the Thing: Ways of Seeing the Teacher *by Ann Cook and Herb Mack*

Teacher Curriculum Work Center: A Descriptive Study *by Sharon Feiman*

Testing and the Testing Industry: A Third View *by John Williams*

Teachers' Seminars on Children's Thinking: A Progress Report *by Bill Hull*

Children's Thinking in the Classroom *by Kathe Jervis*

The Art of Seeing and the Visibility of the Person *by Patricia F. Carini, \$3.50*

Single Copies \$2.00 (unless otherwise indicated) plus 50¢ handling charge for one or more copies; 10 percent of the cost for three or more copies. A check for the proper amount payable to the North Dakota Study Group *must accompany* the order. Mail to: North Dakota Study Group, Box 8039, University of North Dakota, Grand Forks, North Dakota 58202. Copies of this publication are \$3.50 each.

