FOCUS ON EXCEPTIONAL CHILDREN

EVALUATING SPECIAL EDUCATION PROGRAMS

Clifford E. Howe and Marigail E. Fitzgerald[1]

**EVALUATING
SPECIAL EDUCATION
PROGRAMS**

Material for this article has been summarized from work prepared for the Iowa Department of Public Instruction (Howe & Fitzgerald, 1976). It is not intended as the final word in program evaluation, nor as a cookbook set of procedures. It does outline a model which has been field tested, with the most useful aspects retained.

Iowa has recently reorganized into 15 intermediate units called Area Education Agencies (AEA) to provide special education, media, and other services to areas with school populations ranging from a low of approximately 15,000 to one with almost 130,000 pupils. It seems certain that recent legislation will require more accountability and evaluation procedures for individual children than has historically been the case. Iowa's legislation essentially mandates special education for all handicapped students, and requires annual reevaluation of the appropriateness of the program and student placement (State of Iowa, 1974). It is further indicated that school districts, in conjunction with the Area Education Agency, must develop procedures designed to evaluate and improve special education programs and services.

At the federal level, the Education for All Handicapped Children Act (P.L. 94-142, 1975) mandates that each state obtaining funds under this act must provide for evaluation procedures to ensure the effectiveness of programs designed to meet the educational needs of handicapped children (including evaluation of individualized education programs). Evaluation must be carried out at least annually.

Several models have been proposed in the past decade which focus on evaluation in education. One of the most extensive was that developed by Phi Delta Kappa's Research Advisory Committee under the authorship of Stufflebeam and others (1971). Recognizing that evaluation of educational programs was long overdue, they developed a model which combined knowledge of the process as well as product evaluation. Their model views the roles of evaluation as being made up of context, input, process, and product (CIPP). The outcome of evaluation is seen principally as providing useful information for making decisions about program alternatives.

1. Dr. Howe is Professor of Special Education, University of Iowa. Mrs. Fitzgerald is Director of Educational Services, University Hospitals and Clinics, University of Iowa.

Stufflebeam deals with the reasons educational evaluation has either been done poorly or not at all in the past. Included are symptoms such as avoidance (the process is viewed as painful), anxiety (evaluation is viewed as a judgment, often of personal competency), immobilization or lethargy and lack of interest, skepticism regarding whether evaluation can really be done or whether the results are of any use, and a lack of significant differences as the result of much educational research (leading to frustration on the part of the practitioner). Stufflebeam suggests ways in which these difficulties could be overcome and proposes models which will be useful to practitioners (1971).

Popham (1972) proposes a more restricted view of evaluation, and ties it to instructional objectives and criterion-referenced measurement. This handbook provides practical application in the process of constructing and measuring objectives. Emphasis is placed on learner performance data. Evaluation is viewed as a process of determining the desired ends or goals of the educational system, and judging the worth of educational means through both formative (process) and summative (product) assessment. Although his model is not limited to measures of individual student change data, this is the major emphasis.

A third model for educational evaluation is the technique long used by various accrediting associations (NCA, etc.). Stufflebeam (1971) categorizes this as evaluation based on professional judgment. Evaluative criteria are provided, but the major work is done by the schools themselves through a self-study procedure which usually takes about a year. After it is completed, a visiting team of experts and peers comes to the school to observe for a few days, studies the data provided by the school, and renders judgments and recommendations regarding the quality of the program. The major strength of this technique is usually seen in the self-study aspect, where a school and community critically evaluate themselves, thus investing enough of themselves in the evaluation effort to be willing to make and implement decisions for improvement.

An extension of the models above, and one which seems to be very relevant to special education programs, is that of Goal Attainment Scaling (GAS), proposed by Kiresuk and Sherman (1968). These techniques were originally developed from grants by the National Institute of Mental Health, and focused on ways of determining the effectiveness of different treatment approaches for patients in community mental health centers in the Minneapolis area.

In summary, it would appear that successful program evaluation is concerned with two major issues. The first is that of determining the technical approach which appears to have the highest likelihood of yielding useful data to use in making decisions regarding future directions for the program. For the individually-tailored program in special education, it would seem that the use of Goal Attainment Scaling provides promising possibilities.
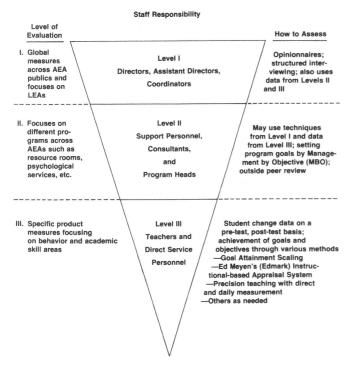
The second issue which seems critical is that of developing a readiness in the organization to undertake program evaluation. Time should be spent with those involved to reduce defensiveness, develop trust, and reach a consensus regarding both the purpose of evaluation and the process to be used. If educational evaluation can be seen, as Stufflebeam (1971) states, "as the process of delineating, obtaining, and providing useful information for judging decision alternatives," then the effort can be viewed in light of its real purpose.

Figure 1 shows that program evaluation must occur throughout the organization and may take different forms at different levels. While evaluation of individual children is specific to each pupil at Level III, it becomes much more general at Level 1. Middle management in Level II uses techniques from both Levels I and III and helps tie the entire process together.

SETTING THE STAGE

For change to occur in an organization, some significant member needs to be the instigator and begin the dialogue with key staff members. Most AEAs in Iowa do not now have an organized program evaluation system, and the director of special education would seem to be the logical person to initiate such an effort. The director must first be convinced of the importance of evaluation efforts. Once

Figure 1
PROPOSED MODEL FOR EVALUATION OF SPECIAL EDUCATION AEAs

Staff Responsibility

Level of Evaluation		How to Assess
I. Global measures across AEA publics and focuses on LEAs	**Level I** Directors, Assistant Directors, Coordinators	Opinionnaires; structured interviewing; also uses data from Levels II and III
II. Focuses on different programs across AEAs such as resource rooms, psychological services, etc.	**Level II** Support Personnel, Consultants, and Program Heads	May use techniques from Level I and data from Level III; setting program goals by Management by Objective (MBO); outside peer review
III. Specific product measures focusing on behavior and academic skill areas	**Level III** Teachers and Direct Service Personnel	Student change data on a pre-test, post-test basis; achievement of goals and objectives through various methods —Goal Attainment Scaling —Ed Meyen's (Edmark) Instructional-based Appraisal System —Precision teaching with direct and daily measurement —Others as needed

this commitment is made by the director, time is needed for informal discussion with the leadership staff.

Our experience suggests that an initial session of not less than two hours is necessary, and that several followup sessions of similar length may be required. During these preliminary discussions, the desired outcome is to overcome resistance to the idea of evaluation and to begin to formulate a plan of positive action. The following topics usually come up in these initial sessions, and need to be resolved:

Q: Are we really going to have to do this sometime in the future in order to get funds, or is this just another passing fad?

A: All signs seem to point in the direction of increasing accountability for appropriateness of programs and to accumulate some evidence showing that the approach used is beneficial. Legislation and Rules in Iowa now require systematic evaluation and will probably become more stringent as the AEAs mature. P.L. 94-142 will become operational in the next year or two, and includes strong statements regarding evaluation plans in order to qualify for these federal funds for the handicapped. It would appear that token evaluation systems won't suffice for the future.

Q: We already have too much to do, and if this becomes another requirement, who is going to do it and where are we going to get the time? How about hiring someone who is a specialist in evaluation?

A: One way to manage an evaluation system is to use "third party" or outside evaluators. This supposedly has the advantage of not adding to the work load of the present staff, and of guaranteeing an objective approach. Many federal projects, such as Title III grants, use

this technique. However, there are disadvantages and one of the major ones is that the existing staffs in an AEA may not invest themselves in the operation and instead see it as a one-time effort conducted by outside paid professionals. Another way to view program evaluation is to think of it as a long-term management technique where the evaluation data become a part of the decision-making process on an ongoing basis. Outside specialists can be used for specific tasks occasionally, but the evaluation effort is really a part of the long-term management of the AEA and all leadership staff should contribute to it.

Q: Aren't we really talking about evaluation of staff, and judgments being made about the personal competence of individuals?

A: It would be dishonest to say that program and staff evaluation are completely separate from each other. Both are important, and need to be done. However, the emphasis on program evaluation is broader than an individual teacher, psychologist, or speech clinician. The purpose is to gain information that is program-wide, such as determining the impact of the total program of speech services in an AEA, for example, rather than focusing on the competency of one speech clinician. Obviously, the success of the total program is dependent on the competency of each person, but program evaluation aggregates the results for the total program, and the individual remains anonymous. Evaluation of an individual serves another purpose and should not be confused with program evaluation.

Q: Many evaluation studies gather dust in the files and much research ends up with conclusions of "no significant differences." How do we know we won't be embarking on a similar enterprise?

A: You don't, really, but that is a strong argument for planning it ourselves as a management tool on a long-term basis, of asking evaluation questions that make sense for our AEA, and of using the results for planning for the next year or several years. We have to make these decisions anyway, and evaluation results should be viewed as just another piece of information which will help us when we must set priorities and make hard decisions.

Q: Okay, I'm convinced. How do I begin with my staff of ten consultants, for whom I am responsible?

A: You need to remember that they will have the same, or perhaps even more, misgivings or questions about evaluation that we've struggled with in the past few hours and sessions. Start slowly, allow at least two hours in a beginning session with all of them, and let them voice all their concerns as well as ideas as to what form the evaluation should take. Expect them to feel avoidance, anxiety, and skepticism, and deal honestly with these feelings. Try to develop credibility of the evaluation process and outline the risks involved. One sure way to encounter difficulty, or perhaps even failure, is to present the evaluation scheme in terms of a dictate from the top. An honest approach might be to say that we will be committed to a continuing evaluation effort, but that the form and methods used are matters to be decided by the group.

The following sections are intended to give ideas and examples of how evaluation might be used by an AEA throughout the various levels of the organization. Evaluation moves from specific child change data as a major component in Level III, to much more general attitudinal data for the total program at Level I.

LEVEL I   EVALUATION PROCEDURES

Level I evaluation is concerned with general views held by consumers throughout the AEA regarding programs

and services for the handicapped. The procedure could be likened to a "Gallup Poll" approach, and samples attitudes from a variety of publics. Stratified random sampling procedures are used so that numbers of respondents are manageable in terms of costs, yet yield a reliable and valid picture of the present status of the organization.

Figure 2 shows an example of an opinionnaire which was field-tested in AEA 16 in the spring of 1976 (Johnson, 1976). Questions included were selected as being of high priority by the leadership staff, including the director and assistant director of special education, supervisors, and consultants.

The staff decided there were eight different subgroups that should be canvassed, including:

Local education agency superintendents
Local education agency building administrators
Special education support staff
Special education instructional staff
Regular education instructional staff
Parent groups for the handicapped
Outside agencies (mental health centers, private schools, etc.)
Secretarial staff in the area education agency

Different subgroups could be chosen by another AEA if, in their judgment, the opinions of such groups were important to the success of the AEA. A minimum sample of 40 was selected from each subgroup with populations which were larger than 40. For those groups whose entire population in the AEA approximated 40 or fewer, all were included in the sample. Where sampling was done, a table of random numbers was used.

A total of 217 opinionnaires was mailed, with 168 returned, for an overall response rate of 77%. This level of response is quite high, considering the diversity of populations sampled and the fact that no additional follow-up mailings were made. The rate of return varied among groups from 100% for superintendents, 85% for building administrators and special education instructional staff, to a low of about 50% for regular teachers and parent groups.

Opinionnaires were returned to a neutral agency outside of the AEA to protect anonymity and encourage more honest responses. Data from the opinionnaire were then tabulated, using a standard computer program which yielded results for each question in the form of means and standard deviations for each of the eight groups. In addition, percentages of each group responding from 5 (agree strongly) to the other extreme of 1 (disagree strongly) were provided on the computer printout for each item.

From this summary, the 18 questions asked were rank-ordered from high to low in terms of degree of agreement. Inspection of the data also pointed up specific subgroups whose responses varied significantly from the overall group.

This information gave the leadership group in AEA 16 a good indication of the general reaction to various aspects of special education services. It is common practice to stop at this point in the procedure and to use these data as input for planning and decision making. However, it is our feeling that another step should be taken to validate results, probe in more depth the indicated problem areas, and solicit suggested changes for improvement. For this followup, we used the technique of structured interviews.

The leadership staff used the opinionnaire data to derive the questions to be asked during the structured interviews, focusing on those areas where there was disagreement among groups or where responses were most negative by the majority of the groups. The advantage of this particular approach is the ability to narrow the range of questions asked and avoid the "shotgun" effect of including many broad questions or of guessing as to what the critical issues are.

Summarizing Structured Interview Data

The interview team should meet with the special education director after interviewing has been completed, to share the major themes which emerged. This should be done while the team is intact and before leaving the AEA. Nothing is more devastating to the process than to have the AEA director wait for several weeks before getting any feedback on results.

After the verbal exit interview with the AEA director and whatever additional staff he or she wishes to include, a short written report is usually prepared within the next week, summarizing strengths and weaknesses observed and indicating possible recommendations. As a final step, and if the special education director so wishes, the chairman of the interview team meets with the director and his or her leadership staff to further discuss the results of the opinionnaire and structured interviewing, and to help plan intervention strategies for future change. At this point, all original data have been returned to the AEA for its use.

Evaluation procedures for Level I are AEA-wide, sample opinions of the various publics served by special education, and combine opinionnaire data with followup structured interviewing. Trust and anonymity are important ingredients for the data to be accurate and useful, and for the process to proceed without undue defensiveness. All original data and results are given to the AEA for its use in assessing the current situation and in planning for the future. If done correctly, the entire process takes a minimum of time and yields useful data.

LEVEL II EVALUATION PROCEDURES

Level II comprises middle management AEA personnel, including supervisors, consultants, and program heads. They are responsible for an area of support services or type of instructional program. Level II personnel also provide a link between instructional programs in the schools and the overall management of special education programs at the director level.

Program evaluation for Level II can utilize the opinionnaire and structured interview techniques outlined earlier under Level I. Questions asked and opinions sought are more narrowly-focused, and related to a specific program such as psychological services, audiology, resource rooms, etc. Because the focus is more delineated, it is possible to go into more depth, as well as to make comparisons of various approaches to delivery of services. Interviewers can be selected from peers in other AEAs throughout the state.

Management by Objective (MBO)

Much has been written about MBO, and the general procedures are common knowledge. We encountered some resistance to the technique and made modifications which retained the major concepts but reduced the amount of detail which is usually associated with the technique. Using an MBO approach can add a major component to program evaluation. It makes planned work efforts explicit, establishes timelines, and pinpoints responsibility. Program development can be planned in a more orderly fashion, reducing the likelihood of decision making occurring as a defensive posture. A staff person's time can be better utilized and focused on those objectives which have higher priority. The technique also provides a concrete basis for staff supervision and for coordination among staff at the middle management level.

Another variant, which we believe has considerable potential, combines elements of MBO and Goal Attainment Scaling. Goal Attainment Scaling techniques are outlined in detail in the next section, but a sample of one Goal Scale is included in Figure 3 (p. 8) to show how major work priorities can be scaled (Howe, 1976).

Note that only major activities are scaled and that the form is not filled with a great amount of detail. All management and supervisory positions involve some maintenance activities which are routine and must be done. However, they should not be the most important activities of a manager, and need not be included in a Goal Scale. Some type of monitoring of a clinical nature will usually suffice to ensure that routine chores are completed on time.

Middle management is a critical part of any organization, and program evaluation is particularly important at Level II positions in AEAs. The master plan has been developed by the state and the AEA director and his or her staff. It is the job of the supervisors, consultants, and program heads to see that the plan is translated into action and to measure the results.

We see middle management as the key to successful program evaluation. It is the supervisor and the consultant who must deal with the problems of efficient application of technical skills to the teaching process, systematic ordering of the process so that teaching can take place with measurable results, and maintenance of teamwork among the principal participants in the process.

LEVEL III EVALUATION PROCEDURES

Evaluation at Level III focuses specifically on child change data. The two most frequent approaches to documenting change in students are the use of pre- and post-test batteries and the recording of progress on specific behavioral objectives. Other useful systems are available commercially, such as the *Instructional Based Appraisal System* (Meyen, 1976) and various remedial curricular programs. Specific behavior recording, precision teaching, and classroom observation approaches may also provide excellent evaluative data on changes in students.

We view evaluation of student progress as primarily a teacher function, since it is most important to the child. The choice of evaluation approach used at Level III depends in part upon the intended use of the results of the evaluation. To meet the intent of recent state and federal legislation for the handicapped, evaluation should provide an annual review of student progress, assist in determining where to go with each child, and should directly relate to the planning and improvement of the instructional program.

An extension of the models cited above, and one which seems to provide promising possibilities to special education programs, is Goal Attainment Scaling as proposed by Kiresuk and Sherman (1968). Goal Attainment Scaling can be viewed as a logical evaluation approach for individualized programs using instructional objectives, prescriptive teaching, and behavior charting methods. It asks the teacher to predict the results for a specified future time and then provides a simple way of scoring the actual outcomes. The method concentrates on the major priorities thought important for each child, and can handle different priorities for different children.

We have piloted Goal Attainment Scaling in a number of communities and with many different types of special education programs. Teachers and support staff have

Figure 2

OPINIONNAIRE

PERCEPTIONS OF SPECIAL EDUCATION SERVICES IN AREA EDUCATION AGENCY 16

This form is an attempt to get feedback from various groups on selected aspects of special education services provided by AEA 16. The purpose is to learn from you those things which you feel the AEA is doing well, and those which should be improved or discontinued. The results will be used to evaluate our present position and to plan for the future.

Below is a series of statements with which you may agree, disagree, or have not had any basis for answering. Please blacken the appropriate box for each statement and return in the enclosed envelope. It takes only 5 or 10 minutes and will be a great help to us.

	Agree strongly	Agree somewhat	Neutral	Disagree somewhat	Disagree strongly	No basis on which to answer
1. More services are being provided for handicapped children as compared to last year	☐	☐	☐	☐	☐	☐
2. Better services are being provided for handicapped children as compared to last year	☐	☐	☐	☐	☐	☐
3. Staff of the AEA are available when I need them: psychologists	☐	☐	☐	☐	☐	☐
consultants	☐	☐	☐	☐	☐	☐
speech clinicians	☐	☐	☐	☐	☐	☐
social workers	☐	☐	☐	☐	☐	☐
special education administrators	☐	☐	☐	☐	☐	☐
hearing services	☐	☐	☐	☐	☐	☐
4. The AEA Division of Special Education should reduce the number of programs it administers directly and turn these over to local school districts	☐	☐	☐	☐	☐	☐
5. Too much money is currently being spent for special education, often at the expense of regular education	☐	☐	☐	☐	☐	☐
6. Decisions made regarding special education at the AEA central office include about the right amount of consultation with those involved at the local level	☐	☐	☐	☐	☐	☐
7. Sufficient supplies and materials are provided for each handicapped child	☐	☐	☐	☐	☐	☐

	Agree strongly	Agree somewhat	Neutral	Disagree somewhat	Disagree stongly	No basis on which to answer
8. The quality of consultation provided by the special education AEA staff is high	☐	☐	☐	☐	☐	☐
9. There is a sufficient quantity of consultation and support staff available from the AEA to adequately support instructional programs for the handicapped	☐	☐	☐	☐	☐	☐
10. The time it takes for a handicapped child to be processed, from original referral to placement, is a good investment of effort	☐	☐	☐	☐	☐	☐
11. Staffing children with various professionals meeting as a group is a good idea	☐	☐	☐	☐	☐	☐
12. The system of referral for a child is working well	☐	☐	☐	☐	☐	☐
13. A major reason for AEA providing some services is the lack of program offerings by local school districts	☐	☐	☐	☐	☐	☐

ADD ANY STATEMENTS THAT YOU THINK SHOULD HAVE BEEN INCLUDED:

14.	☐	☐	☐	☐	☐	☐
15.	☐	☐	☐	☐	☐	☐
16.	☐	☐	☐	☐	☐	☐

OTHER COMMENTS YOU MAY WISH TO MAKE:

Figure 3

GOAL ATTAINMENT SCALE

August 1976	June 1977		Student Name
Start Date	**Score Date**		

	Howe		School
	Teacher		

Score	Percentile		Town
	*Entry Level	× Exit Level	

SCALE HEADINGS:	SCALE 1: MONITORING AEA 16	SCALE 2: NEW PROGRAM DEVELOPMENT	SCALE 3: TRAINING STATE CONSULTANTS	SCALE 4: STATE PLAN
LEVELS:	(weight 1 =)	(weight 2 =)	(weight 3 =)	(weight 4 =)
Most unfavorable outcome thought likely	AEA 16 will abandon current plan and will not invest further in program evaluation efforts.	No new AEAs will participate in program evaluation efforts beyond token involvement.	DPI staff fails to become involved and does not attend any inservice training sessions.	AEAs request more staff and money with no evidence as to effectiveness of current operation.
Less than expected success	No further refinement of current evaluation plan.	One or two new AEAs will develop segments of an evaluation system.	Minimum of two DPI staff will participate in one or more inservice training sessions.	Majority of requests for staff and funds based on subjective opinion and emotion.
Expected level of success	AEA 16 will continue evaluation effort and operationalize pilot plans for collecting child change data with monitoring by me.	Minimum of three new AEAs will develop a program evaluation system with me in 1976-77.	Frank Vance, John Lanhan and two consultants will participate with me in inservice training of AEA staff for evaluation.	Evidence of some "outcome data" in state plan and of use of such data in future planning decisions.
More than expected success	Complete AEA-wide evaluation system will be operational with monitoring by me.	Four or five AEAs will participate in developing an evaluation system.	Several DPI staff will assume leadership in developing and monitoring AEA evaluation systems.	Annual plan and fiscal requests tied to evaluation evidence of current operation.
Most favorable outcome thought likely	Complete AEA-wide evaluation system will be operational and function without my help or monitoring.	More than five AEAs will develop a system.	State Division of Special Education will require an evaluation system of AEAs and will monitor progress.	Additional positions authorized to AEAs by State Superintendent based on outcome data showing success of program.

been largely enthusiastic in learning the technique, and have found it a satisfying approach to individualizing program efforts and in evaluating their outcomes. Mastery of the technique comes from actually sitting down

and writing scales for individual children. The first ones will be laborious and time-consuming. A complete Goal Attainment Scale can be written in about one-half hour after having done the first five or ten.

Goal Attainment Scaling Procedures

Scale Development. A number of priority areas should be selected for the student. These priorities will not necessarily include all the important work to be done with each child, but should be representative of the major goal areas to be concentrated upon in the special education program during the time covered by the Goal Attainment Scale. Typically, these major problem areas will have been identified in the child's staffing. Goals can then be determined by the teacher and support staff charged with responsibility for planning the child's program.

Once the priority areas for scaling have been identified, each should be given a title. The title may be abstract, theoretical, or vague. It is designed to focus the attention of someone inspecting the scale on the major goal areas being evaluated. The title may also be thought of as the place where the teacher constructing the scale has an opportunity to indicate the general problem area to which the specific variables described in the body of the scale correspond.

When priority areas have been selected and titles identified for the scale, a numerical weight (numbers 1, 2, 3, 4, or 5) can be added to each scale below the title. The weighting system indicates the relative importance of the scale. The scales can also be used without weighting if all goals are judged to be equally important. The higher the number used, the more significant the scale is relative to other scales. The title box can also be used to indicate any special sources of information for the scale, such as normative test data like the KeyMath or Durrell Reading Tests. Figure 4 shows a complete Goal Attainment Scale example where the specific priority areas are not equally important. The scales have been weighted 5, 4, 3, and 4 respectively (Leone, 1975).

The key level for predictive purposes is the expected level, or middle box, on each five-point scale. The expected level presents the best, most realistic prediction possible of the outcome which will have been attained by the student at the score date. The statements ought to be realistic, so that the expected level of each scale reflects what outcome realistically could be attained by the score date, not necessarily what should be attained. The estimate of the expected outcome ought to be independent of the student's current level of functioning. It may be that the expected level outcome would reflect no change or even regression; in spite of the undesirableness of this situation, it belongs in the middle box if this outcome is thought most likely.

The expected level is usually developed first and should be the most likely outcome. The other outcome levels should be constructed after the expected level and

should be thought less likely to occur. It is not required that all levels be written in on the scale, but at least one box on each side of the middle box must be specified. Thus, at least three of the five boxes or levels must be completed for each scale.

The "more than expected success" and "most favorable outcome thought likely" levels offer teaching objectives and guide program efforts and planning in the future. Although humanitarian instincts would lead us to hope we could accomplish these higher outcomes, the accurate use of the Goal Attainment Scaling technique would not allow these levels to be reached very frequently. Similarly, the "less than expected success" and the "most unfavorable outcome thought likely" should not occur as frequently as the middle box outcome. Nevertheless, these less favorable outcomes are important to balance the picture of possible outcomes, to pinpoint children and priority areas needing closer evaluation, and to help judge when special needs go beyond the program's capacity to meet them.

We recommend that more than one person be involved in writing the child's Goal Attainment Scale. In Iowa, the child's teacher, the Area Education Agency special education consultant, and the program's supervisor may share this responsibility and periodically meet and confer on the child's progress and the program's usefulness. The team may use this format to clarify and differentiate responsibilities in accomplishing the predicted outcomes.

Having a team of professionals involved in the Goal Attainment Scaling process provides a check and balance mechanism to avoid setting expected level statements unrealistically high or low. If the outcome statements must be agreed to by the program supervisor and the special education consultant, the likelihood of setting expected levels too low to "look good" or too high is minimized. With the use of pre/post normative test data as an additional element in the total evaluation process, the concern and/or likelihood of such an event occurring is reduced.

Learning to write Goal Scales is a developmental process. You learn primarily by experience, and gradually improve in being able to specify level outcomes within priority areas. Although an individual program is written for each student, there is some overlap of scales among students. After having written a number of scales, you will find that you can often draw from the bank of earlier scales for some items.

Scoring and Interpretation. The student's level of functioning at the time the scale is developed can be noted on the Goal Scale form by placing an asterisk in each box for entry level. At the followup score date, the scales are marked with an "X" for outcome. Two possible kinds of

Figure 4

GOAL ATTAINMENT SCALE

9-1-75	6-1-75		
Start Date	**Score Date**		**Student Name**

	P. Leone		**School**
	Teacher		

Score	**Percentile**		**Town**
	*Entry Level		**× Exit Level**

SCALE HEADINGS:	**COURSE CREDITS**	**TASK BEHAVIOR**	**USE OF LEISURE TIME**	**REGULAR CLASS INTEGRATION**
LEVELS:	(weight 1 = 5)	(weight 2 = 4)	(weight 3 = 3)	(weight 4 = 4)
Most unfavorable outcome thought likely	Will fail all 7th grade classes and/or be excluded from continuing attendance in regular classes.	Almost no work accomplished in spite of continual teacher intervention.	Continual pattern of starting and soon thereafter quitting.	More than 20 hours per week of special class contact.
Less than expected success	Will fail two regular 7th grade classes.	Completes a quarter to a half of assignments with continual prodding and ultimatums.	Completes most of requirements with frequent adult supervision (3 or more per week).	12-20 hours of special class contact.
Expected level of success	Will fail one regular 7th grade class, but will pass on to 8th grade.	Completes half of expected assignments in resource room with frequent prodding.	Joins sports team, league, or club and completes season with periodic adult supervision (once per week).	Maintains appropriate behavior to the extent that he will have only 11 special class contact hours each week during last quarter.
More than expected success	Will successfully pass all 7th grade classes with D or C average.	Completes most assignments with 1 or 2 reminders.	Continues in 2 or more after-school groups with periodic supervision.	1-10 hours of special class contact.
Most favorable outcome thought likely	Will successfully pass all 7th grade classes with a B average.	Completes most assignments with no reminders.	Completes requirements of one team, club, league, or group with no special supervision.	Maintained completely as regular 7th grade student with no direct special class service.

effectiveness measures can be collected from the Goal Attainment Scaling system: Whether or not the "expected" levels of outcome are reached, and whether or not change occurred. The degree of change can also be documented on the basis of the post-test data or records gathered at the predetermined date. To score the Goal Attainment Scale, scores of −2 (most unfavorable), −1 (less than expected), 0 (expected level,) +1 (more than expected), and +2 (most favorable) are given for each final outcome. A formula or calculation table is then used to convert these scores to standard scores with a mean of 50 and a standard deviation of 10.

In order to add composite Goal Attainment Scores of various children or to compare one child longitudinally, some cautions of a statistical nature should be kept in mind. The various goal scales should be done realistically, so that the expected mean value for the group is near zero (standard score = 50) and with a standard deviation approximating one (converted standard score = 10). Stated another way, about two-thirds of the composite scores of the total group should fall within the "expected level of success" on the Goal Attainment Scale; about 10% to 15% should obtain scores of "less than expected success"; 10% to 15% should receive "more than expected success"; and very few (2% to 5%) should achieve scores at the extremes of "most favorable" or "most unfavorable" outcome thought likely. If there is considerable variance from this as a group, the effect is that some composite goal scores will have heavier weights than others and make the comparisons less valid. The cautions just noted should not discourage use of the technique, but should be kept in mind.

Applications for Use. As indicated earlier, results from individual evaluation scales are potentially useful in a variety of applications. First, scored scales provide information on individual child changes upon which decisions regarding the student's continuation or change in placement can be initiated and resolved. Secondly, the scored scale in and of itself is a data base on the child's placement that can be inserted into the student's cumulative folder for documentation of placement and subsequent instruction. Thirdly, the scale provides the teacher, the Area Education Agency special education consultant, and the program's supervisor a systematic means to review the achievements of the students served in the program.

The potential for comparing program models as well as types of instructional designs (behavior modifications, use of paraprofessionals and/or associates in instruction, teaching methods, etc.) over the years is an aspect yet to be explored as the implementation of the process proceeds to its conclusion. Indeed, the use of Goal Attainment Scaling as described in this proposed model appears to be most promising.

REFERENCES

Howe, C. E. *Goal scaling used to evaluate major work priorities.* Unpublished manuscript, College of Education, Division of Special Education, University of Iowa, Iowa City, IA, 1976.

Howe, C. E., & Fitzgerald, M. E. *A model for evaluating special education programs in Iowa Area Education Agencies.* Des Moines, IA: Iowa State Department of Public Instruction, Division of Special Education, 1976.

Johnson, W. *An example of an opinionnaire field tested in AEA 16 in the spring of 1976.* Unpublished manuscript, Director of Special Education, Area Education Agency 16, Mount Pleasant, IA, 1976.

Kiresuk, T., & Sherman, R. Goal attainment scaling: A general method for evaluating comprehensive mental health programs. *Community Mental Health Journal,* 1968, *4*(6), 443-453.

Leone, P. *Sample of complete goal attainment scale example where specific priority areas are not equally important.* Unpublished manuscript, Child Psychiatry Service, University of Iowa Hospitals and Clinics, Iowa City, IA, 1975.

Meyen, E. *Instructional based appraisal system.* Bellevue, WA: Edmark Associates, 1976.

Popham, W. J. *An evaluation guidebook: A set of practical guidelines for the educational evaluator.* Los Angeles: The Instructional Objectives Exchange, 1972.

State of Iowa. *Rules of special education.* Des Moines: Department of Public Instruction, 1974.

Stufflebeam, D. L., et al. *Educational evaluation and decision making.* Itasca, IL: Peacock Publishers, 1971.

ADDITIONAL REFERENCES

Garwick, G. *Guidelines for goal attainment scaling.* (A product of the Program Evaluation Resource Center, T. J. Kiresuk, Director.) Grant No. 1 R12 MH1561902. Washington, D.C.: National Institute of Mental Health, Department of Health, Education, and Welfare, February 1975.

Kiresuk, T. J., & Garwick, G. *Program evaluation project report 1969-73: Basic goal attainment scaling procedures* (2nd ed.). Grant No. 5 R01 1678904 and Grant No. 1 R12 MH2561902. Washington, D.C.: National Institute of Mental Health, Department of Health, Education, and Welfare, 1975.

Neufeldt, A. H. Considerations in the implementation of program evaluations. In Davidson, Clark, & Hamerlynch (Eds.), *Evaluation of behavioral programs.* Champaign, IL: Research Press, 1973, pp. 65-82.

Senf, G., & Anderson, D. *A program evaluation manual for project initiators.* Final Report Project No. H 12-7145B, Grant No. OEG-0-714425. Washington, D.C.: Bureau of Education for the Handicapped, June 1974.

ALERT

March 9-12, 1977
Association for Children with Learning Disabilities
Washington Hilton Hotel
Washington, D.C.

March 15-19, 1977
National Association of School Psychologists
Cincinnati, Ohio

April 11-15, 1977
Council for Exceptional Children
International Convention
Atlanta, Georgia

May 8-10, 1977
Canadian Speech and Hearing Association
Empress Hotel
Victoria, British Columbia, Canada

INDEX   VOLUME 8

AUTHOR INDEX

Armstrong, Jenny R. (December 1976)

Beck, G. Robin (April 1976)

Belmore, Jane (May 1976)

Belmore, Ken (May 1976)

Capps, Lelon R. (January 1977)

Fitzgerald, Marigail E. (February 1977)

Haring, Norris G. (April 1976)

Hatfield, Mary M. (January 1977)

Hayden, Alice H. (April 1976)

Holloway, Alwyn H. (March, April, September, October, 1976)

Howe, Clifford E. (February 1977)

Jens, Ken G. (May 1976)

Keogh, Barbara K. (March 1976)

Kohfeldt, Joyce (October 1976)

Levitt, Marc L. (March 1976)

Moran, Mary Ross (November 1976)

Morse, William C. (September 1976)

CHRONOLOGICAL INDEX OF TITLES

Special education in the mainstream: A confrontation of limitations? (March 1976)

General principles and guidelines in "programming" for severely handicapped children and young adults (April 1976)

Language programming for the severely handicapped (May 1976)

The helping teacher/crisis teacher concept (September 1976)

Blueprints for construction—teacher-made or teacher-adapted materials (October 1976)

The teacher's role in referral for testing and interpretation of reports (November 1976)

Individually guided education (IGE): One model for mainstreaming (December 1976)

Mathematical concepts and skills—Diagnosis, prescription and correction of deficiencies (January 1977)

Evaluating special education programs (February 1977)

CLASSROOM FORUM (March, April, September, October, November, 1976)