

# PROGRAM EVALUATION

*There are two ways in which evaluation information might be used: To prove or to improve*

(C M Callahan and M Caldwell, 1995)

### WHY MUST PROGRAMS BE EVALUATED?

Developers of programs for the gifted tend not to evaluate the success of their programs or the effectiveness of program components. There seem to be several reasons for this reluctance. First, they typically feel that because they created their program in good faith it necessarily is "successful." Second, they prefer to invest their time in planning and teaching. "The primacy must be the intervention, not the evaluation," according to Cook (1986). A third reason is that "success" in teaching gifted and talented students is difficult to assess, compared with using achievement test data to evaluate remedial or basic skills programs. A fourth reason for hesitancy is that the evaluation results could threaten the program itself. For example, if the creativity tests, thinking skills tests, self-concept inventories, or other measures of complex constructs fail to show an improvement—perhaps because the instruments were unreliable or unrelated to what was taught—the data might be taken as evidence of program failure.

Traxler (1987) surveyed 192 school districts having gifted programs. She found that half of the programs were not evaluated at all, and of those evaluated most did not employ trained evaluators. Observations by teachers and evaluations of student products were the most common types of assessments.

Although gifted programs are more difficult to evaluate than other programs, this evaluation is vital. Gifted programs come and go; the record of continuity is dismal. Therefore, if teachers and program directors hope to maintain or expand their programs,

they must be able to demonstrate the success of the program to their administration, to school board members, to parents, and to state or federal funding sources. This is *accountability*. These publics will want to know who is being served by the program, how they are being served, and the beneficial effects of the program. They also will want to know if the program is cost-effective. Equally important, teachers and program directors need information that will allow them to revise and improve the program. Beyond creating classroom quizzes or evaluating student papers and projects, teachers and coordinators usually have little training or experience in educational evaluation. This chapter is intended to simplify and clarify the evaluation of G/T programs and guide the teacher or coordinator in the evaluation process.

### EVALUATION DESIGN: BEGIN AT THE BEGINNING

Although evaluation is the topic of the last chapter in this book, evaluation of a gifted program belongs at the *beginning* of program planning. At the outset, when setting goals and objectives for a G/T program, one should design a methodology for measuring whether or not those objectives are reached (e.g., Tomlinson and Callahan, 1994).

### "Difficult" and "Easy" Evaluations

Callahan (1986; see also Callahan, 1993) set forth some of the difficulties that are unique to evalua-

tions of gifted programs. She pointed out that (1) "no agreed upon standards of good programming exist within the field of gifted education and (2) many of the objectives in programs for the gifted are very complex and not easily defined"

Some examples of *difficult* objectives for evaluation are improvements in leadership, self-awareness, self-concept, decision making, reasoning, analyzing, synthesizing, evaluating, social responsibility, intrinsic motivation, critical thinking, and creative thinking. Other objectives are comparatively *easy* to evaluate. Acceleration programs, for example, provide almost self-evident evaluation data. Did students succeed in the advanced classes, the college courses, or the correspondence courses? Did the grade-skipping or the early admission to kindergarten work well for the students involved? Enrichment plans that result in a bona fide product—a school newspaper, a report of a research project, a poetry book, a dramatic production, artwork, a movie—also provide relatively easy evaluation data. Such products reflect a clear change in student skills and performances that, most likely, would not have occurred without the G/T program (Renzulli and Smith, 1979). Smith, LeRose, and Clasen (1991) used easily-obtained statistics to show a profound effect of the Racine, Wisconsin, Lighthouse Project (LeRose, 1977, 1978): None of the minority students who participated in the program dropped out of high school and 76 percent planned to attend college. Among comparably talented minority students in the Racine district, 45 percent dropped out.

Whether objectives are difficult or easy to evaluate, we nonetheless must try to evaluate every planned objective.

## EVALUATION MODELS

There are many models for structuring the evaluation of education programs. Several will be summarized here in an admittedly oversimplified fashion. In all cases, the intrigued reader will need to explore the more complete, original statements.

Provus' (1972) *discrepancy* model assumes five stages in the creation of a program. At each

stage the reality of the program is compared with a standard, and any discrepancy is corrected. In Stage 1, *Design*, the initial program plan is compared with a set of theory-based design criteria, perhaps as defined by an outside consultant. If there is a discrepancy, the program plan is modified accordingly. In Stage 2, *Installation*, the reality of the program as it is implemented is compared with the design adopted in Stage 1. Again, any discrepancies between program design and installation will guide changes. These changes could be in the installation or in the Stage 1 design criteria. In Stage 3, *Process*, the actual program activities are compared with the proposed program activities, and any discrepancies will result in corrective alternatives. Stage 3 is especially important for creating an effective, successful program. In Stage 4, *Product*, actual student products are compared with the planned ones. This may be the main evaluation of program objectives. Stage 5, *Product Comparison*, involves a comparison of students' products and learning outcomes with those of other programs to determine program efficiency in the cost-benefit sense.

Part of Eash's (1972; see also Renzulli, 1975) *differential evaluation* model involves three considerations in the evaluation process: (1) *Effort*—how time is spent (that is, the program activities); (2) *Effect*—products and outcomes; and (3) *Efficiency*—the relationship of effort and resources to the quality of effects realized. As the program evolves from a newly planned, innovative program (initiatory stage), through the implementation and field testing of the program (developmental stage), to the established and stable level (integrated stage), more emphasis is placed on evaluating effects and efficiency (2 and 3).

The Renzulli and Ward (1969) DESDEG (*Diagnostic and Evaluative Scales for Differential Education for the Gifted*) model was designed specifically for evaluating programs for gifted and talented students. It also may be used for program planning and development (Renzulli, 1975). DESDEG includes a set of five published documents corresponding to the five parts of the model. Part I is the Manual, which explains every-

thing you ever wanted to know about DESDEG Part II, Evaluative Scales, includes scales for evaluating each of 15 “ideally conceived educational practices” or Program Requirements, which subdivide into the five Key Features (general areas) in Table 18.1. Part III consists of Basic Information Forms, “a comprehensive inventory of factual information about all aspects of a program . . . organized and keyed . . . to each of the five Key Features. . . .” These aid in the objective collection of data. Part IV is the Evaluator’s Workbook, designed to aid the evaluator in handling the information derived from the Basic Information Forms and from observations. Part V is the Summary Report, which (1) permits the evaluator to transfer numerical data to statistical and graphic summary sheets, and (2) aids in the creation of a summary narrative related to each Program Requirement.

It can be seen even in these sketchy outlines that the evaluation of education programs, including G/T programs, can be approached from many

different viewpoints, use different strategies, and focus on a variety of dimensions and considerations.

### THE RIMM MODEL

Rimm’s (1977) model both (1) structures program evaluation in a relatively easy-to-follow fashion and (2) ties it to the initial program plan. Summarizing a program in one picture is very helpful for conceptualizing program components and therefore for relating evaluation needs to those components. Figure 18.1 demonstrates how the different parts of a program fit together and, importantly, how evaluation can help one monitor all educational *Inputs (resources)*, all *Processes (activities)*, and all *Outcomes (goals and objectives)*. Using the model has many advantages. First, it helps us understand the relationships among educational resources, processes, and outcomes. Second, using the model helps prevent the

**TABLE 18.1** The DESDEG Model

Key Feature A: Philosophy and Objectives	
Program Requirement 1:	Existence and Adequacy of a Document
Program Requirement 2:	Application of the Document
Key Feature B: Student Identification and Placement	
Program Requirement 3:	Validity of Conception and Adequacy of Procedures
Program Requirement 4:	Appropriateness of Relationship Between Capacity and Curriculum
Key Feature C: The Curriculum	
Program Requirement 5:	Relevance of Conception
Program Requirement 6:	Comprehensiveness
Program Requirement 7:	Articulation
Program Requirement 8:	Adequacy of Instructional Facilities
Key Feature D: The Teacher	
Program Requirement 9:	Selection
Program Requirement 10:	Training
Key Feature E: Program Organization and Operation	
Program Requirement 11:	General Staff Orientation
Program Requirement 12:	Administrative Responsibility and Leadership
Program Requirement 13:	Functional Adequacy of the Organization
Program Requirement 14:	Financial Allocation
Program Requirement 15:	Provision for Evaluation

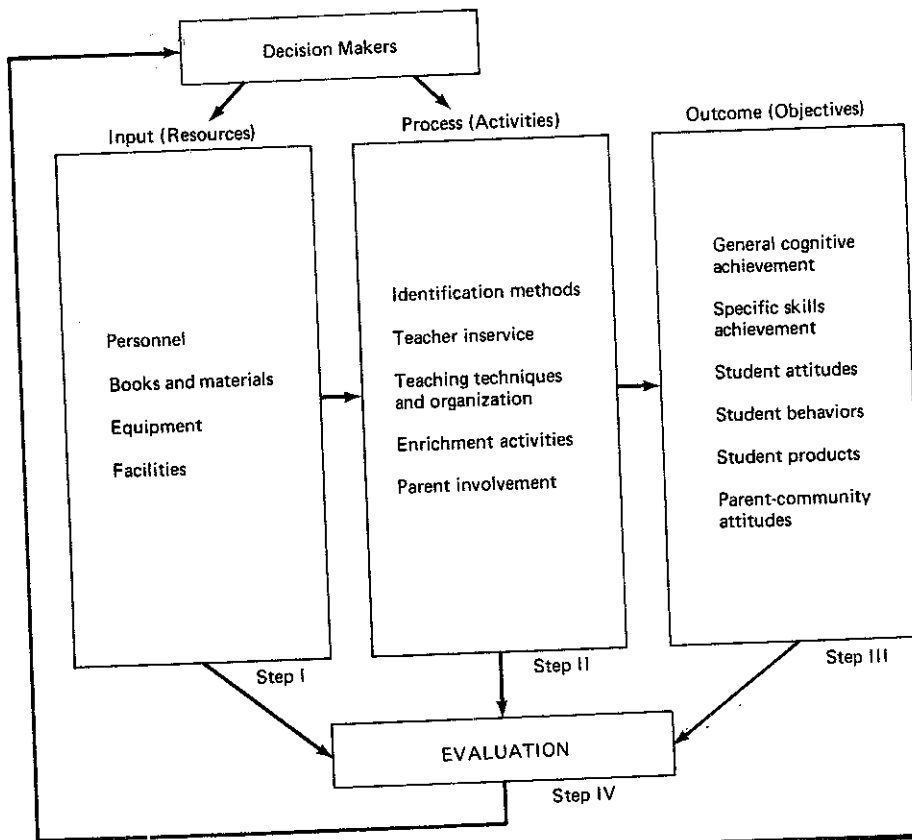


FIGURE 18.1 Framework for the Evaluation and Monitoring of a Gifted Program.

implementation of any activity without considering its eventual *evaluation*. Third, the model helps us become more sensitive to the close relationship of program decisions to the many student outcomes. Finally, and most importantly, the model itemizes on one page the program components that should be evaluated in regard to both (1) how well the component was implemented and/or (2) how successfully that component helped achieve program goals.

The components within each of the three steps of the model reflect specific areas that should be evaluated. That is, each of the various types of input/resources (Step I), processes/activities (Step II), and outcomes/objectives (Step III) should be evaluated. Evaluation data (Step IV) from the components will present a comprehensive picture of

the success and impact of a gifted program. This information is brought together and fed back to the decision makers who will use it for further planning—for modifying the input and process steps, which may include program expansion. Without the crucial evaluation step, there would be little clear basis for good decisions. The success of this approach depends heavily on the relevance and the clarity of the evaluation information obtained in Step IV.

The reader may wish to draw a blank framework to outline his or her own program and evaluation model and complete the diagram as this chapter is read.

Step I: *Input* represents resources. Resources typically include such program ingredients as teaching and support personnel, books, materials,

equipment, and facilities. Resources also may include more specific categories such as community resource persons, specific student populations, or funding sources. Resources are the investments in the program and they usually are relatively easy to identify and list

Step II: *Process* includes the activities of the program—everything that is planned to make the program effective. Typical categories of activities include identification procedures, teaching techniques, educational groupings, enrichment experiences, acceleration plans, teacher in-service training, and parent involvement activities. One may wish to itemize more specific curriculum activity components of the program, for example, creative-thinking instruction, creative writing, a computer mini-course, accelerated mathematics, Renzulli's (1977) Enrichment Triad model, and so on. The list of categories of activities should be exhaustive, which means that each and every specific program activity would fit into a process category

Step III: *Outcome* represents the goals and objectives of the gifted program. It actually may be easier to complete the list for Step III *before* completing Steps I and II. What do you expect to accomplish? What are the purposes of the program? Ideas for program goals were listed in Chapter 3.

Note that increased academic achievement is not necessarily a central outcome of a gifted program, although it may be an objective for gifted underachievers or, for example, for an accelerated math, biology, or reading program. More frequently, increases in specific *skills* are the intended outcomes of a program, for example, creative, critical, or evaluative thinking and independent study and research skills. Positive student *attitudes*, including self-concepts, attitudes toward education, and high career aspirations, also are frequently stated goals. In addition, scientific, literary, and artistic *products* are important and potentially measurable objectives. Parent and community attitudes toward the program and toward the needs of gifted children are further important outcomes to monitor.

Completing the model by inserting program components into each of the four steps should provide the reader with an overview of the entire program and assist with planning the necessary evaluations.

### COMPLEXITY OF EVALUATION AND AUDIENCE: A HIERARCHY

The sophistication of an evaluation will be related to the intended audience—the persons who are the decision makers for a particular program. These decision makers can be placed in a *hierarchy* in terms of the quantity and the quality (statistical complexity) of the information they must have to carry out their own responsibilities. The goal should be to provide the appropriate information to match the information needs of these decision makers. Callahan and Caldwell (1995) refer to audiences as *stakeholders*. A suggested form for identifying relevant stakeholders can be found in Appendix 18.1.

Briefly, (1) students and parents will need less information than (2) teachers and program directors, who in turn need less information than (3) administrators and school board members (4). The state department of education will require still more technical information, and (5) the federal government, with its highly experienced grant reviewers, will require the greatest amount and the highest technical level of information.

Table 18.2 summarizes the relationship between levels of the decision-making hierarchy, components of the program emphasized, and the persons primarily responsible for the evaluation.

The hierarchy of decision makers is based on the different purposes or uses of the program evaluation information. A student or his or her parents may need to know only if the activity is generally interesting, challenging, motivating, and beneficial in very personal terms in order to decide if the student should continue in the program. A relatively small amount of information is needed for a decision that may have an important impact upon just one student. The student and his or her parents, with the help of the program teacher and/or

**TABLE 18.2** Summary of the Relationship Between Hierarchy of Decision Making and Evaluation Plan

<i>Hierarchy of Decision Making</i>	<i>Model Component Emphasized</i>	<i>Staff Responsibility</i>
Individual Student or Parent	Individual Process and Outcome	Student, Teacher, and/or Parent
Teacher or Program Director	Process and Outcomes	Teacher, Program Director, or Administrator with some background in evaluation or Independent Evaluator
State or Federal Government	Input, Process, and Outcomes	Independent Evaluator with support of Program Staff

coordinator, can assess the value of the program for that student

The teacher or program director requires quite a bit more information in order to modify, improve, and perhaps expand the program. This function of evaluation is called *formative evaluation*. Conducted throughout the school year, it is intended to provide immediate and continuous feedback to the staff regarding program strengths and weaknesses. The main focus of formative evaluation will be on *process* (Figure 18.1); that is, on the value of various activities and experiences. House and Lapan (1994) recommend that more formative evaluation be done because it requires less formalistic design and serves to improve the program. Formative evaluation usually is conducted quite effectively by program staff, although observations of an independent (outside) evaluator usually add objectivity, insights, and ideas to the evaluation.

The school principal, district administrators, and members of the elected school board must decide whether to continue, change, or expand the gifted program, and hopefully not to fold it up. To make these kinds of decisions, a *summative evaluation* must be conducted. The emphasis is on *outcomes*, and so a summative evaluation is conducted at the end of a unit, project, or most often,

the school year. It “sums up” program success. Administrators and school board members may expect the staff to relate program *input* and *process* to outcomes in order to estimate some kind of cost-effectiveness. A teacher or program director with training in evaluation and statistics sometimes can successfully conduct such an evaluation. Board of education members primarily will want to know if a program has been “effective,” and so the *hows* and *whys* of its effectiveness must be clearly and simply communicated.

Generally, local school people will feel their accountability obligation is met if they can prove that (1) the program was conducted as planned, (2) the students learned what was taught, and (3) the process of learning was a positive one.

However, if a program is state or federally funded, a professional summative evaluation conducted by an experienced outside evaluator probably will be necessary. This evaluation typically includes a more technical, experimentally oriented evaluation. The teacher’s role then becomes one of *cooperator*. The teacher need only inform the evaluator of the resources, activities, and objectives, and provide him or her with the needed data. Some test administration may be involved. Clear communication with the evaluator and full cooperation in gathering the necessary data will facilitate an

accurate evaluation of the program. The outside evaluator should not be considered an adversary who is anxious to pounce on weaknesses. His or her role is to provide constructive feedback to the staff for program improvement, and to objectively report the reasonableness and effectiveness of the program plans, methods, activities, and so forth. Although the objectivity of outside evaluators may appear threatening during early program development, that same objectivity will be extraordinarily reinforcing when a program is stabilized and functioning well.

The outside evaluation may require the use of a *control group* of subjects (for example, students of similar ability in another school district where there is no program). Test scores from such a control group, when compared with scores of students in the program, would help determine if any improvements (for instance, in creativity scores, self-concept development, or achievement) are due to the program and not due simply to maturation, the passage of time, or other educational experiences. Teacher-coordinators likely will drown in a sea of statistics if they attempt this type of evaluation on their own.

As emphasized earlier, an important first consideration for effective evaluation is outlining the evaluation design at the same time the rest of the program is planned, namely, before the program begins. Even an expert evaluator will be less able to do his or her job if the program begins without coordinating the program objectives with their eventual evaluation. Beginning a G/T program without an evaluation design is comparable to beginning one's classroom teaching without a curriculum plan. In both cases, the preplanning helps outline where you are going, how to get there, and what to do when you arrive.

In summary, no program for the gifted should be conducted without some evaluation. Whether the decision makers are students, parents, teachers, administrators, school board members, or state or federal agencies, they will want to know about the "success" and the particular "effects" of the program.

## INSTRUMENT SELECTION

Some form of measurement is almost always necessary to determine the degree to which program objectives have been achieved (e.g., those in the *outcome* rectangle of Figure 18.1). Ideally, in order to reach sound conclusions one should try to obtain *two* measurements of each objective, particularly the most important ones. Also, whenever possible one should use instruments—tests, questionnaires, rating forms—already available. To do a proper job of developing one's own instruments requires a considerable amount of time and usually requires training and experience in test construction.

Besides, with a little digging a teacher-coordinator most likely will discover that the test-building work already has been done. Many instruments for assessing innumerable aspects of G/T programs are available, for example, in Renzulli (1975), Renzulli, Reis, and Smith (1981), Renzulli and Reis (1985), Sjogren, Hopkins, and Gooler (1975), and the Association for the Gifted Evaluation Committee (1979). If the same or similar program plans are used in several schools or districts, evaluation questionnaires may be shared.

A parent questionnaire for an Enrichment Triad Model, Revolving Door Model, or Schoolwide Enrichment Model program appears in Appendix 18.2 (Renzulli and Reis, 1985). Teacher, student, and parent questionnaires are available for the evaluation of many components of the Renzulli program models, for example, a *Student Product Assessment Form*, *Type I Enrichment Evaluation Form*, *Scale for Evaluating Creativity Teaching Materials*, and others (see Renzulli and Reis, 1985).

For new evaluation ideas, check current journals on gifted education and creativity and documents of the National Research Center on the Gifted and Talented (e.g., Delcourt, Lloyd, Cornell, and Goldberg, 1994); they frequently describe instruments that have been recently developed and validated. Some sample instruments appear at the end of this chapter.

Although there are indeed many different tests for review, it can happen that an instrument exactly right for a specific purpose will not be found. Do not make the mistake of using an instrument—no matter how carefully designed—that measures something other than what the program plans to teach. It is essential that the *objectives* of the program activities, on one hand, be matched with the *purposes* and the *contents* of the tests and inventories, on the other. If test purposes or contents do not match program objectives and activities, one is not very likely to find a measurable effect of the program. This simple point may seem self-evident to the sensible reader. However, it is a common error for G/T teachers and coordinators to teach one set of contents and skills (for example, creativity), yet evaluate others (for example, advanced reading)—and then be surprised and disappointed to find “no effects” or “no transfer” of the training experience. Standardized tests may be especially inappropriate in evaluating disadvantaged gifted programs. Test scores for these programs may reflect socioeconomic status, linguistic abilities, or cultural factors rather than the varied dimensions of the actual gifted program (House and Lapan, 1994). As mentioned earlier, negative evaluation findings can easily be misused: They can readily and logically be interpreted as evidence that the program was ineffective.

### Pilot Testing

A program evaluator may need to *pilot* a test, that is, try it out with a few children to help decide if it is appropriate for the desired purpose. For example, suppose a program included accelerated reading or math and the evaluator wanted to determine if a particular norm-referenced test could be used to evaluate student achievement.<sup>1</sup> If the test is administered to two or three students and they “top

out” near the maximum possible score, this would immediately signal that the test does not assess the high achievement level of the gifted students. It may be necessary to pilot three or four tests to find an appropriate one. One strategy is to pilot several tests simultaneously with several small groups of students. This quickly provides plenty of information for test comparison and selection.

### Topping Out, Regression Toward the Mean, and Reliability

As discussed earlier in this book, the topping-out phenomenon is a frequent occurrence when gifted students take standardized, norm-referenced achievement tests. The tests often are not difficult enough to discriminate among or evaluate the learning of high-achieving gifted children. If we measure improvements from pretests to post-tests but students already have achieved ceiling scores on the pretests, it will appear that the students showed no measurable improvement. In fact, there simply was no place for the scores to go. Actual progress could not be measured.

A potentially more serious problem related to using too-easy achievement tests comes from the *regression toward the mean* effect. This simply means that given a first score that is extreme (either very high or very low), by chance alone the next score is likely to “regress” toward the mean. Now if some gifted students score near the very top of a pretest, their scores on the post-test by chance alone are likely to be lower—incorrectly suggesting that participation in the special G/T activities damaged their learning. With regressed scores on a basic skills test, an audience could conclude that the gifted program is having a negative effect on the basic skills of students. Because pulling students out of the regular classroom often is a controversial issue in the first place, such an interpretation could be devastating to an excellent program.

Tests should be precise enough to measure changes from pretest to post-test. Therefore, tests should have good test-retest and/or alternate

<sup>1</sup>A *norm-referenced* test is one of moderate difficulty that is designed to produce a normal, bell-curve distribution of scores. For example, a standardized achievement test



forms reliability. Rating scales must have a sufficiently wide scoring range to detect pre- to post-test differences (Beggs, Mouw, and Barton, 1989).

Divergent thinking tests can produce different results depending on variations in instructions, differences in motivational levels, or other reasons. In one evaluation by the second author, post-test scores on the *Torrance Tests of Creative Thinking* (TTCT) were significantly lower than pretest scores, despite an excellent creative thinking program. Perhaps the lower scores were due to a less motivating classroom climate toward the end of the year, compared with the more stimulating "new opportunity" at the beginning of the program. It also was possible that the two TTCT test forms are not equivalent (i.e., the more difficult form was given as a post-test); or that these divergent thinking tests simply did not measure the kinds of creative dispositions and abilities that were learned (Davis, 1992). Whatever the reason(s), such findings can be damaging to a good program.

### TEST CONSTRUCTION

There are times when a state or federal agency will require evidence for the effectiveness of a program, but the exactly appropriate tests and measures simply do not exist. It therefore may be necessary to construct original tests or questionnaires in order to evaluate specific student skills, information, attitudes, or abilities; the effectiveness of specific program components; or even the overall quality of the total program. Technical help from an evaluation expert at a university or private consulting firm, someone experienced in evaluating G/T programs and in constructing tests, probably will be needed.

To select a consultant, contact directors of other G/T programs and state or even national leaders in gifted education to find who in the area is available and qualified. When a consultant is recommended, one can elicit opinions of his or her work from teachers and program directors for whom the consultant has provided services. One also can ask to review tests and reports the consultant has prepared for other clients.

Competitive bidding can be misleading, since there is a considerable range of quality that fixed evaluation dollars can buy. Also, a poor instrument or evaluation is worse than none at all, even if the cost is low. Test construction may be expensive, which is one good reason to use established tests, as recommended earlier.

### Rating Students' Products

Artistic, literary, scientific, or other types of student products may be outcomes of a program, and although they are difficult to evaluate reliably, their quality is measurable. Product evaluation usually involves either of two approaches. With the first *gain score* approach, samples of students' work obtained at the outset of the program (pretest products) are compared with students' products at the end of the educational experience (post-test products). Gain scores may be used to evaluate products that reflect the development of skills and abilities in art, creative writing, divergent thinking, technical skills, or other areas.

With the second *absolute* approach, individual students' products are evaluated according to their excellence without an objective comparison with earlier products. The absolute approach would be used if science or other major projects (for example, of the Renzulli Type III variety) are evaluated for which earlier comparison products are not available.

With either the gain score or the absolute method, rating scales may be used to quantify the judgments. As two examples, Renzulli (1975) recommended the scale in Appendix 18.3, developed by the Warwick, Rhode Island, Public Schools, for evaluating the quality of visual arts products. The scale in Appendix 4.5 may be used to evaluate the excellence of virtually any type of student product.

The rating scales in Appendices 18.3 and 4.5 could serve as models for creating one's own product-rating scales. Alternatively, the following steps could be used to develop and apply original rating scales:

**Step 1. Determine the evaluation criteria (for example, creativeness, technical skill) and the specific scales for rating those criteria.** The criteria should come from the teaching objectives. Note the objectives reflected in each scale in Appendices 18.3 and 4.5. As for the rating scales themselves, while Appendix 18.3 uses six-point scales, a five-point scale as in Appendix 4.5 usually is most comfortable. It is best to describe the meaning of each point on the scale. The “low,” “moderate,” and “high” descriptions of Appendix 18.3 and the “To a great extent,” “Somewhat,” and “To a limited extent” of Appendix 4.5 are very minimal descriptions. For example, a scale for evaluating the use of *humor* in creative writing might be based on the following descriptors:

- “1”: No use of humor
- “2”: One humorous comment, which does not appear original.
- “3”: Two or three humorous statements or paragraphs, which do not appear original.
- “4”: Two or three statements or paragraphs that do appear original.
- “5”: Original humorous themes skillfully integrated into the entire story.

Such scale-point descriptions will improve the accuracy (reliability) of the ratings.

**Step 2. Collect pretest products from each child in the program.** All products should be identified with code numbers so that the students’ names and the date of collection will not be obvious to the raters.

**Step 3. Select at least two raters.** They may be teachers or community members. Train them to use the particular scales with extra products that will not be used as evaluation data. After rating several products together, use a few more extra products for individual raters each to rate separately.

**Step 4. Calculate the percentage of rater agreement.** With five-point scales, two raters should agree on at least 80 percent of the practice ratings

before they begin rating the pretest and post-test products. If the raters do not agree at an 80 percent level, they need more training, including a discussion of the reasons for their disagreements.

**Step 5. After interrater reliability is established, to minimize bias, raters should rate all products.** This should be done without knowing the names of the children and without knowing which are pretest products and which are post-test products.

**Step 6. Calculate the average pretest ratings for the group and compare them with the average post-test ratings.** If a program is effective, the average post-test rating should be higher, indicating that skill development increased since the pretest products were created. Statistics may be necessary in order to conclude that the overall increase is not due to chance. Also, as noted earlier, a control group may be necessary to prove that the improvement of the trained students was not due to the passage of time or to other educational experiences.

For information regarding the gains (or losses) of individual students, one would examine the differences between pretest and post-test ratings for each student.

**Step 7. In reporting the findings, the means (and other statistics, if any) must be meaningful, reliable, and valid.** If possible, include a few sample products (for example, samples of creative writing, scientific reports, or photos of artwork) to illustrate the student gains. These sample products will help any audience understand the meaning of the ratings.

Hennessey and Amabile (1988) used a much more informal system for rating the creativeness of children’s stories given orally in response to pictures. Three teachers, experienced in children’s creative writing, were simply asked to rate the creativeness of each story “using your own, subjective definition of creativity.” They were given no training and made their assessments independently. Despite the lack of training and simplicity of the rating task, interjudge reliability was a high

0.91. These results are sufficiently impressive and surprising that your authors would recommend a replication before using teacher ratings of creative stories, for example, as a prudent measure of the success of a creative writing program.

### Classroom Observation Data

Parents, administrators, school board members, and government agencies usually like to know what happens in a gifted program, and so classroom observation data are good data to collect. A structured observation form can be developed for a program and used to describe "who is doing what with whom and when" in a very specific way. As one example, Appendix 18.4 (from Rimm, 1981) shows a structured observation form used for monitoring a reading program. The letters at the top of each column represent each of the ten students observed. Filling in a circle indicates that a student is involved in one of the activities listed in the rows. An objective observer would enter the classroom at random times of the day and observe ten randomly selected students, recording their activities on the checklist. Descriptive comment may be added to each observation to provide a richer description of the class environment.

A form such as the one in Appendix 18.4 could be adapted for use with any particular G/T program. Category headings could be modified or new ones created; each would have specific subcategories. For example, one might wish to record student's interest level, student's behavior, student's interactions, teacher's instructional activities, materials and equipment in use, or activities of other adults. Additional headings, each with specific subcategories, could be included depending upon the special activities of the program.

If desirable, at the end of the year summary forms can be tabulated and percentages calculated to describe the specific *type* and *extent* of the activities engaged in during that year. This summary provides objective documentation of the year's program activities. Thus if a school board member wanted to know how much time microcomputers were in use, how many students partic-

ipated in the *Great Books* discussions, or what proportion of the G/T program time was spent in independent projects, numbers and percentages would be available to support personal observations and impressions. Nothing is quite so convincing as hard data

### Questionnaires

Activities in both the *process* and *outcomes* components of Rimm's model may be evaluated with questionnaires. The best way to find answers is to ask questions. Decision makers at all levels will want to know (and have a right to know) the effectiveness and special strengths of a program as well as its weaknesses as perceived by others. They also are interested in others' constructive suggestions for improving the program. If questionnaires are brief and require only that a few numbers be circled and/or a few questions be answered, most persons will respond. You can bet that those who are strongly enthusiastic and those who are most disappointed or critical will be certain to respond.

Questionnaires use various types of items. Objective items include *checklists*, *rating scales*, *rankings*, and *multiple-choice* statements. The advantages of objective items include ease of development, efficiency of administration, clear response options, easy and objective scoring, and ready quantifiability for statistical purposes. Some disadvantages include the limited nature of the response options, along with little or no information about the *reasons* for the judgments.

Because of these disadvantages, many objective questionnaires also include open-ended items. Open-ended items provide an opportunity for students, teachers, parents, principals, and/or school board members to voice reasons for their opinions, as well as to contribute suggestions and potential solutions to problems. Information can indeed be rich and valuable. For formative evaluation purposes, open-ended items usually are more valuable than objective items. In the negative column, open-ended items are more time-consuming for both respondents and scorers, interpretations may be

ambiguous, and the answers are not easily quantified.

As examples of combining objective with open-ended questionnaires, Appendix 18.5 presents a brief questionnaire that could be used by teacher participants in an in-service program. Appendix 18.6 shows a questionnaire that allows school board members to evaluate G/T services of a high school. Appendix 18.7 presents a non-objective open-ended form that students can use to evaluate a learning center.

Generally, it is wise to set aside a few minutes in every program for teachers and students to complete the evaluation forms. If you create the time, persons will fill out the questionnaires on the spot; questionnaires taken home or mailed often are not returned.

### Interviews

Interviews can be used to gather information about input, process, or outcomes. Fetterman (1993) described several kinds of interviews. Formally structured and semistructured interviews are similar to questionnaires in that they use specific formats and questions, and so data can be compared. Informal interviews are easiest to conduct and often are spontaneous, almost as if they were extensions of conversation. Fetterman points out that informal interviews may answer some important questions that would not be answered by more formal interviews.

Retrospective interviews involve reconstructing the past. Although not the most accurate, they not only are useful but sometimes are the only type of interview that can provide insight into difficulties. Retrospective interviews can be either formal or informal.

### DAILY LOGS

As a general principle, *log everything*. Each staff member should keep a notebook handy in which to make brief notes on daily activities. The value of the entries will far outweigh the few minutes invested each day. The kinds of information that could be logged include:

- A description of activities
- Preparation for the activities
- Number of participants
- Perceived effectiveness of the activities
- Modifications for the future
- Any data collected
- Any anecdotal material

A personal log kept by each staff member can provide important program documentation serving at least three purposes. First and foremost, it can provide a description of the activities and accomplishments of the students and, therefore, of the value of the specific learning activities and projects. Second, it can assure administration and board members that the staff member has indeed made critical contributions to the program. Third, it will serve to remind the teacher of the quantity and quality of his or her own contributions and accomplishments.

### INDICATORS

Measures including high-school dropout rates, increased attendance, and percentages of students going to college are recommended by House and Lapan (1994) to evaluate gifted programs for disadvantaged students. Other indicators, such as number of students taking advanced courses or classes in which they will continue after programming, could be helpful for some gifted programs.

Indicators should be tied specifically to the objectives of the gifted program. For example, a program targeted toward encouraging gifted girls to achieve should compare indicators of female success to past female performance rather than to male performance. Hopefully, the gender gap in performance will eventually close, but female success should not become a contest between the sexes. Improvement of female performance is the most important indicator of an effective program for females.

Although it is difficult for schools to do follow-up studies of high school graduates, data on academic adjustment to college, percentage of high school students who complete college, attendance at graduate school, and career choices all

could be valuable indicators of successful gifted programs. House and Lapan (1994) point out that studies typically are funded for short periods of time, and demands for accountability tend to be short term. Nevertheless, whenever longitudinal studies are possible, they provide important evaluation information.

### STUDENT SELF-EVALUATIONS

Student self-evaluations are important in G/T programs. Primarily, they provide individual students with a clear measure of accomplishment relative to their own goals and objectives. Positive feedback reinforces student motivation and commitment, while objectively documenting their personal progress. For program evaluators, when individual self-evaluations are combined they become important measures of student outcomes, some of which could not be obtained in any other way.

Independent self-monitoring usually can be conducted quite handily by students in junior and senior high school. Younger children also can take responsibility for self-evaluation, with a little help from their teachers. One example of a summative, end-of-the-year student self-evaluation form appears in Appendix 18.8.

### PERFORMANCE CONTRACTING

Student *performance contracting* is another vehicle for individual student evaluation. Within the student contract, the teacher and the student spell out:

1. Specific *objectives*, including skills to be learned and final projects and papers to be completed.
2. The *activities* in which the student plans to engage to achieve the objectives.
3. The *deadline* by which the objectives will be completed.
4. The *materials* the student will produce (or collect) in support of his or her attainment of the objectives.

5. The *methods* and *criteria* by which the attainment of the objectives will be evaluated.

The student contract is a "study guide" for both the student and the teacher, as well as the basis for an evaluation and documentation of the student's personal performance. Performance contracts are a type of learning activity that encourages the independence and creativity that most gifted students thrive on. Thus individual student needs are served while also providing program evaluation and accountability data.

### QUALITATIVE AND QUANTITATIVE EVALUATION

Although some evaluation techniques have already been described, Linnemeyer's (1994) list of evaluation techniques in Table 18.3 can help program coordinators remember some of the many qualitative and quantitative techniques that are available.

### COMMITMENT TO EVALUATION

Teachers and coordinators in gifted programs are likely to view evaluation as burdensome. Their time always is scarce, and time used for evaluation will be at the expense of time for students. It thus is very tempting simply to avoid evaluation altogether. However, skipping the evaluation process is a short-sighted decision for gifted programs—which, we repeat, have a history of being quickly cut from district, state, and federal budgets. Good evaluation is the only way to determine the most effective way to enhance the education of gifted learners. It also is the only way to prove to sponsors and decision makers that the program is indeed accomplishing its objectives.

As a final note, Table 18.4 summarizes guidelines proposed by House and Lapan (1994) for evaluating programs for gifted disadvantaged students. These same guidelines seem very appropriate for evaluating other gifted programs.

TABLE 18.3

<i>Qualitative Evaluation</i>	
Case Studies	Logs
Inventories	Autobiographies
Questionnaires	Portfolios
Scrapbooks	Videos
Recordings	Tape Recordings
Report Cards (Descriptive)	Diaries
Behavior Journals	Activity Records
Collections	Anecdotal Records
Samples of Work	Profiles
Writing Samples	Scope and Sequence Charts
Sociograms	Student Interest Surveys
Self-Concept Scales	Essays
<i>Quantitative Evaluation</i>	
Achievement Tests	Mental Ability Tests
Checklists	Creativity Tests
Criterion Referenced Tests	Auditions
Pupil Graphs	Grades
Critical Thinking tests	Music Tests
Report Cards (Numerical)	Teacher-Made Tests
Charts	Classroom Climate Tests
Sociometric Tests	Art Tests
Leadership Tests	Psychomotor Ability Tests

From Linnemeyer (1994), by permission of the Illinois Association for Gifted Children.

## SUMMARY

Evaluation in gifted education has been minimal. However, it is important both for demonstrating success to outsiders and for improving the program. Evaluation plans should be made at the outset of program planning.

Some objectives are difficult to evaluate, such as improvements in self-awareness, creativity, analyzing, and social responsibility. Other outcomes are comparatively easy to assess, such as the success of acceleration or the improved quality of student products.

Provus' discrepancy model includes five steps: Design, Installation, Process (activities), Product, and Product Comparison. At each step one compares the program reality with a standard and then

corrects the discrepancy. Eash's differential evaluation model focuses on three considerations, Effort (activities), Effect (products and outcomes), and Efficiency (the relationships of effort and resources to the quality of the Effects).

Renzulli's DESDEQ model included five steps corresponding to five documents: the Manual, Evaluation Scales, Basic Information Forms, Evaluator's Workbook, and the Summary Report.

Rimm's model structures the program evaluation and ties it to the initial program plan. The three steps of Input (resources), Process (activities), and Outcome (objectives), each with specific subcategories, all may be evaluated.

Increased academic achievement may or may not be a central outcome. Rather, improvements

TABLE 18.4

*Guidelines for Conducting an Evaluation of Programs for the Gifted Disadvantaged*

- Students are identified by (nontraditional) procedures that are appropriate to the particular types of students and the program.
- Critical components of the program have been adequately described in the evaluation, including descriptions of the actual classroom teaching.
- Higher level thought processes, creativity, and other particularly appropriate outcomes for gifted students, depending on the particular program, have been examined by appropriate means other than standardized tests.
- Multiple measures have been employed.
- Possible negative side-effects of the program have been examined.
- Other unanticipated effects on students, teachers, parents, the community have been examined.
- The program contains culturally relevant material for culturally different students and recognizes and respects their particular cultural identity.
- The program displays sensitivity to females and includes materials with female role models, etc.

From House and Lapan (1994), by permission of the Association for the Gifted

in process skills and attitudes are usual goals of G/T programs.

Audiences form a hierarchy in the quality and quantity of needed evaluation information. Students and parents need relatively little information to decide whether to continue in the program. Teachers and program directors require more information, particularly continuous, formative evaluation for program improvement. Administrators and school board members will require summative information to decide whether to continue or expand the program. State or federal funding sources will need considerable detailed information, including, for example, test scores and statistical comparisons with control groups.

One should try to obtain two measurements of each important objective. It usually is easier and cheaper to locate already validated tests than to construct your own. It is important to be certain that the test measures the objectives that were the basis of the teaching. Using the wrong tests will produce negative results, creating a bad impression. Standardized tests may be inappropriate for evaluating gifted programs for the disadvantaged.

Pilot testing is advisable, for example, to cope

with the topping-out problem common with gifted students. On a second testing, very high scores may regress toward the mean, creating the appearance that the program damaged, for example, basic skill development. Tests should be sufficiently precise (reliable, wide scoring range) to measure pre- to post-test changes.

With a gain-score approach, ratings of preprogram projects are compared with ratings of post-program projects. With the absolute approach, complex projects are evaluated without comparison to earlier projects.

In creating original rating scales, one would determine the evaluation criteria, collect pretest products, select and train at least two raters, determine rater agreement, use "blind" ratings, and compare average pretest ratings with average post-test ratings. It is desirable to include sample projects in final reports.

Classroom observation data present objective information regarding "what happens" in a gifted program.

The process (activities) and outcome (objectives) sections of Rimm's model may be evaluated with questionnaires.

Objective questionnaire items (for example, checklists, rating scales, multiple-choice questions) are easily administered and scored. They may be combined with less objective and more time-consuming—but highly informative—open-ended questions.

Interviews are useful for gathering information about input, process, and outcomes. Informal interviews are easier to conduct and may yield more information than formal interviews. Retrospective interviews, which reconstruct the past, can provide insights.

Daily logs provide valuable records of activities, preparation, participants, perceived effectiveness, ideas for modifications, and anecdotal information.

The use of indicators (e.g., dropout rates) are recommended for the evaluation of gifted pro-

grams. Indicators should be tied to the objectives of specific programs, such as those for disadvantaged and female gifted students. Follow-up studies, although difficult to conduct, also provide valuable indicators of a program's success.

Self-evaluations provide positive, motivating feedback to students as well as unique program evaluation data.

Performance contracting can be used to individualize instruction and to document student accomplishments.

There are many qualitative and quantitative evaluation techniques available to gifted program coordinators.

Good evaluation is absolutely essential for the continuity and improvement of any gifted program.



**APPENDIX 18.1 IDENTIFYING RELEVANT STAKEHOLDERS**

We have started a list of relevant stakeholders for you, but this list is *only a beginning*.

1. Read over this list and *eliminate* any individuals or groups that are not important to your program.
2. Read over the list and *add* any individuals or groups that are significant to your program.

***School Administrators***

Superintendent \_\_\_\_\_  
Associate Superintendent for Instruction \_\_\_\_\_  
Principals \_\_\_\_\_  
Supervisors or Directors of Instruction \_\_\_\_\_

***Program Staff***

Program Coordinator/Director \_\_\_\_\_  
Teachers in the Program \_\_\_\_\_  
Mentors \_\_\_\_\_

***Other Decision-Makers***

School Board Members \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

***Other School Staff***

Regular Classroom Teachers (if they are not program staff) \_\_\_\_\_  
Librarians \_\_\_\_\_  
School Psychologists \_\_\_\_\_  
Counselors \_\_\_\_\_

***Other Community Members***

President of the PTA \_\_\_\_\_

***Parents of Gifted Students***

***Parents of Students not in the Program***

***Students in the Program***

***Funding Agents***

Source: *A Practitioner's Guide to Evaluating Programs for the Gifted*, a service publication of the National Association for Gifted Children, by C. M. Callahan and M. S. Caldwell, University of Virginia, 1995

## APPENDIX 18.2 SAMPLE PAGE FROM A PARENT QUESTIONNAIRE

**Revolving Door TAG Program Parent Questionnaire****Directions:**

Please do not sign your name to this questionnaire. No attempt will be made to identify persons completing these forms. Please return the questionnaire in the enclosed envelope within the next two or three days.

You can help to make our program better by giving careful thought to each of the questions that follow. Because of the relatively small number of persons involved in the project, each person's opinion will weigh heavily in analyzing the results. We appreciate your cooperation and assistance in helping us to evaluate this program. We suggest that you read through the entire questionnaire before you begin to answer the questions.

- |  | YES | NO  |
|--|-----|-----|
| 1. Did you attend the meeting held earlier this year at which we explained the Revolving Door TAG Program?   | ___ | ___ |
| 2. Do you feel that you have been provided with enough information about the Revolving Door Identification Model in general?   | ___ | ___ |
| 3. Are you familiar with the procedures for revolving a student into the resource room?  | ___ | ___ |
| 4. Has your child been revolved into the resource room this year?  | ___ | ___ |
| 5. If the answer to Question No. 4 is yes, are you familiar with the project or topic that your child worked on while in the resource room?  | ___ | ___ |
| 6. Are you familiar with the procedures for revolving a student out of the resource room?  | ___ | ___ |
| 7. Are you familiar with a procedure we use in the TAG Program called "Curriculum Compacting"?   | ___ | ___ |
| 8. Have you been invited to visit the resource room?   | ___ | ___ |
| 9. Do you feel that you have been offered sufficient opportunity to contact the school or make an appointment to discuss your child's work in the TAG Program?                         | ___ | ___ |
| 10. Has your child encountered any problems with his/her friends as a result of being involved in the TAG Program? If yes, please explain.   | ___ | ___ |
| 11. Has your child encountered any problems as a result of leaving the regular classroom to participate in the TAG resource room? If yes, please explain.                              | ___ | ___ |
| 12. Has your child expressed a concern about missing work in the regular classroom or making up assignments because he or she is out of the classroom to attend the TAG resource room? | ___ | ___ |
| 13. Have you had any communication with your child's classroom teacher about his/her participation in the TAG Program? If yes, please describe.  | ___ | ___ |

**APPENDIX 18.3 PROJECT GIFTED: EVALUATION SCALE FOR VISUAL ARTS**

Code No. _____	Age _____	Boy/Girl _____	TOTAL SCORE _____			
Evaluator _____		Position _____		Date _____		
Elements	Low		Moderate		High	
	1	2	3	4	5	6
1. Creative Expression, Imagination, Uniqueness						
2. Flexibility, Appreciation, and Adaptability to Various Media						
3. Fluency, Variety or Number of Ideas						
4. Sensitivity-Composition-Design						
5. Manipulative Skills: Construction, Weaving, etc.						
6. Growth						
Column Total						
Weight	1	2	3	4	5	6
Weighted Column Total						
TOTAL SCORE						

J. S. Renzulli, *An Evaluation of Project Gifted*. Storrs, University of Connecticut, 1973. Reprinted by permission

APPENDIX 18.4 EXAMPLE OF A STRUCTURED OBSERVATION FORM

Date	a.m.	p.m.	Observer	abc	def	ghi	abc	def	ghi
1. <u>Pupil's Location</u>									
Desk or table	000	000	0000	000	000	0000	000	000	0000
Carrel	000	000	0000	000	000	0000	000	000	0000
Open area	000	000	0000	000	000	0000	000	000	0000
Materials center	000	000	0000	000	000	0000	000	000	0000
Other (specify):	000	000	0000	000	000	0000	000	000	0000
2. <u>Instructional Content</u>									
Readiness	000	000	0000	000	000	0000	000	000	0000
Decoding skills	000	000	0000	000	000	0000	000	000	0000
Comprehension	000	000	0000	000	000	0000	000	000	0000
Enjoyment or appreciation	000	000	0000	000	000	0000	000	000	0000
Vocabulary	000	000	0000	000	000	0000	000	000	0000
Spelling	000	000	0000	000	000	0000	000	000	0000
Grammar	000	000	0000	000	000	0000	000	000	0000
Composition	000	000	0000	000	000	0000	000	000	0000
Oral expression	000	000	0000	000	000	0000	000	000	0000
School library usage	000	000	0000	000	000	0000	000	000	0000
Speed reading	000	000	0000	000	000	0000	000	000	0000
Dictionary skills	000	000	0000	000	000	0000	000	000	0000
Other (specify):	000	000	0000	000	000	0000	000	000	0000
3. <u>Pupil's Instructional Grouping</u>									
Whole-class group instruction	000	000	0000	000	000	0000	000	000	0000
Partial-class group instruction	000	000	0000	000	000	0000	000	000	0000
Tutorial (one to one) instruction	000	000	0000	000	000	0000	000	000	0000
Independent work on individually-assigned activity	000	000	0000	000	000	0000	000	000	0000
Independent work on a group-assigned activity	000	000	0000	000	000	0000	000	000	0000
Shared work on group-assigned activity	000	000	0000	000	000	0000	000	000	0000
Shifting group patterns	000	000	0000	000	000	0000	000	000	0000
4. <u>Instructional and Audiovisual Materials and Equipment in Use</u>									
5. <u>Pupil's Behavior</u>									
6. <u>Person Relating to Pupil</u>									
7. <u>Person's Instructional Role</u>									

S. Rimm, "Evaluation of gifted programs—as easy as ABC." In R. E. Clasen et al. (eds.), *Programming for the Gifted, Talented, and Creative* (Madison: University of Wisconsin, 1981).

APPENDIX 18.5 EVALUATION FORM FOR IN-SERVICE WORKSHOPS

Workshop Title: CREATIVITY IN THE CLASSROOM

Circle the appropriate numbers below.

1. I found this program to be . . . . .

1	2	3	4	5
Dull		Of average Interest		Very Interesting

2. I think that what I learned today will be . . . . .

1	2	3	4	5
Useless		Somewhat Useful		Very Useful

3. I would like to have more inservice programs on this topic.

1	2	3	4
No, not at all	Yes, but not for a while	Yes, more this year	Other (Explain Below):

Explanation: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

4. The things I liked most about this in-service were: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

5. The things I liked least about this in-service were: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## APPENDIX 18.6 SCHOOL BOARD QUESTIONNAIRE

**APPENDIX 18.6 SCHOOL BOARD QUESTIONNAIRE**

We suggest you read the questionnaire before you start to answer.

Please indicate the extent to which you agree or disagree with each of the following statements by circling the appropriate letter. The letters mean the following:

SA — Strongly agree  
 A — Generally agree  
 U — Undecided  
 D — Generally disagree  
 SD — Strongly disagree

Please use the comment line if you want to explain your answer.  
 Answer question 9 only if you have children in the high school.

- |    |   |    |   |   |   |    |
|----|---|----|---|---|---|----|
| 1. | School E High School provides a well-balanced educational program.                                      | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 2. | This school has a good program for able students.   | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 3. | Most parents feel School E is a good school.  | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 4. | The School E program is mainly for the college-bound student.   | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 5. | Students at School E are receiving a good education in the basics like math, English, history, science. | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 6. | There are too many frills in the School E program.  | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 7. | School E has a good extracurricular program, i.e., athletics, music, school paper, drama, clubs, etc.   | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 8. | School E should have more vocational courses.   | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |
| 9. | Most of the teachers our child has had at School E have done a good job.                                | SA | A | U | D | SD |
|    | Comment _____   |    |   |   |   |    |

APPENDIX 18.7 STUDENT EVALUATION: LEARNING CENTER PROGRAM

STUDENT EVALUATION  
LEARNING CENTER PROGRAM

For the past sixteen weeks you have been attending sessions at the \_\_\_\_\_ County Learning Center. We would like to know some of your feelings about the program. By answering questions and completing the following sentences, you can help us in improving the program.

1. Which class did you like best? \_\_\_\_\_
2. Why? \_\_\_\_\_
3. Of the classes I was not in, I wish I could have taken \_\_\_\_\_  
\_\_\_\_\_
4. Why? \_\_\_\_\_
5. I wish my classes at the Learning Center were longer \_\_\_\_\_, shorter \_\_\_\_\_, the same \_\_\_\_\_ (check one).
6. The Learning Center needs more \_\_\_\_\_
7. The class in which I learned or accomplished most was \_\_\_\_\_  
\_\_\_\_\_
8. If I could change three things about the Learning Center, I would
  - a. \_\_\_\_\_
  - b. \_\_\_\_\_
  - c. \_\_\_\_\_
9. Has the Learning Center helped you in any way with things you do at school? \_\_\_\_\_
10. How? \_\_\_\_\_
11. Has the Learning Center helped you in any way with things you do at home? \_\_\_\_\_
12. How? \_\_\_\_\_
13. Has the Learning Center helped in any way with the way you get along with or feel about people? \_\_\_\_\_ If so, how? \_\_\_\_\_  
\_\_\_\_\_

From Florida's *State Resource Manual for gifted Child Education*. State of Florida Department of Education, 1973. Reprinted by permission

## APPENDIX 18.8 PUPIL SELF-EVALUATION

## PUPIL SELF-EVALUATION

Pupil's Name \_\_\_\_\_

Think of yourself at the present time in comparison to last year. As a result of this year's work, please rate yourself on the following items. Place the letters a, b, c, d, and e on the line following each item according to the scale below.

- |                    |               |
|--------------------|---------------|
| (a) Much less      | (d) More      |
| (b) Less           | (e) Much more |
| (c) About the same |               |

1. Ability to think things through for yourself \_\_\_\_\_
2. Knowledge of subject matter areas (science, social studies, and others I have taken) \_\_\_\_\_
3. Interest in school \_\_\_\_\_
4. Ability to see how things go together in a situation (see relationships) \_\_\_\_\_
5. Ability to find information \_\_\_\_\_
6. Ability to work well by myself \_\_\_\_\_
7. The liking and respect of other pupils for me \_\_\_\_\_
8. Ability to judge the usefulness of facts \_\_\_\_\_
9. Ability to get along with my teacher(s) \_\_\_\_\_
10. Enjoyment of learning \_\_\_\_\_
11. Knowledge of arithmetic, spelling, and other basic skills \_\_\_\_\_
12. Curiosity about learning new things \_\_\_\_\_
13. Ability to accept responsibility \_\_\_\_\_
14. Opportunity to make things, experiment, and use ideas \_\_\_\_\_
15. Knowledge of my strengths and weaknesses \_\_\_\_\_
16. Willingness to do work as a leader \_\_\_\_\_

Please answer the following questions:

17. Has the school year been helpful to you? Yes \_\_\_\_\_ No \_\_\_\_\_  
Please explain.

\_\_\_\_\_

18. Has any part of the school work this year created any problems for you? Yes \_\_\_\_\_ No \_\_\_\_\_  
Please explain.

\_\_\_\_\_