

Renzulli, J. S. (1967, March 26-April 1). *Designing an Instrument for Evaluating Programs for Differential Education for the Gifted* [Selected convention paper]. 45th Annual Convention of the Council for Exceptional Children, St. Louis, Missouri, United States.

Designing an Instrument for Evaluating Programs for Differential Education for the Gifted

Joseph S. Renzulli

Introduction

The difficulty of validly appraising the effectiveness of educational programs poses a serious and longstanding problem for those persons who have undertaken the responsibility of educating the nation's youth. In the area of education for the gifted, the problem of formal program evaluation has been compounded by the absence of appropriate means explicitly designed to evaluate the particularized objectives that guide and direct the learning experiences of gifted individuals.

The literature on the gifted reveals a striking contrast between a nearly universal plea for program evaluation and a near dearth of analytical and procedural studies attempting to bring this problem into manageable proportions. In developing the present Instrument, an attempt was made to fill the significant gap that exists between what is widely called for as desirable (i.e., program evaluation) and feasible procedures through which the evaluation of programs of differential education for the gifted (DEG) can be achieved.

Logic Underlying Design of the Instrument

The purpose of the Instrument is to provide a practical means by which educational programs for the gifted may be evaluated. In view of the absence of external criteria capable of distinguishing between varying degrees of program quality, an attempt was made to develop an instrument based on the reflective judgment of a small but eminently qualified group of insightful educators who are students of the subject of differential education for the gifted.

This rationale is based on the belief that research into problem areas for which the significant variables and processes have not been readily defined and made subject to control must first satisfy and be guided by forms of inquiry that are essentially rational in nature. Rationality in the form of human judgment is always an acceptable beginning for the process of structuring a problem area for empirical attack and is a necessary procedure for devising particular methods and techniques for scientific research. Although empirical procedures have been utilized in certain phases of the present endeavor, initial and central dependence on the validity of the instrument is based on the judgment of persons who are considered to be representative of the best thinking presently available in the field of education for the gifted. The opinions derived from a small but select group of persons who have demonstrated an exceptional degree of interest and achievement in this field are considered to be substantially more consequential than opinions gathered en masse from a larger but minimally involved group. The

statistical advantages that may be derived from large scale polling do not suffice in a design which demands that content validity be represented in the initial input. Qualitative judgment, systematically and selectively procured, is manifestly a more promising basis for program evaluation than the wholesale solicitation of opinion from persons who have not been totally immersed in the problem.

Procedures in Constructing the Instrument

The procedures employed in constructing and judgmentally validating the experimental instrument consisted of the following interrelated steps.

First, the literature dealing with program evaluation in various areas of education was searched in order to identify the principal aspects of the problem and to locate relevant information and ideas that might prove useful in developing the present instrument. This initial step also included a nationwide survey which was aimed at locating all existing instruments in use at state and local levels for purposes of evaluating special programs for the gifted. Information resulting from this search was selectively utilized in carrying out certain of the procedures that follow.

The second step involved the selection of a panel of 24 expert judges. This task was accomplished by first identifying, according to a number of specified criteria, a larger group of persons who have made certain contributions to the field of DEG, and by asking this group to nominate, from among themselves, those persons whom they considered to be the most qualified for judging the adequacy of educational experiences for superior and talented students.

Procedurally, the third step consisted of isolating a number of basic characteristics or key features of programs for the gifted. Key features may be thought of as a basic core of minimally essential elements that represent the best practices of mature and excellent programs designed to meet the needs of gifted individuals. The concept of key features represents an essential part of the rationale upon which the Ward-Renzulli scales are based. Out of the entire span of characteristics upon which any educational program might possibly be evaluated—from the quality of the classroom teacher to the adequacy of the supplies and materials, certain program characteristics are manifestly more consequential than others. With respect to the whole array of practices and provisions that possess potential although varying degrees of value in furthering the objectives of DEG, the concept of key features holds that the evaluation of a minimal number of highly significant features will facilitate the evaluative process. The rationale is that, if the more essential features of a program are found to be present and operating excellently, then the probability of less significant and critical features being similarly present is good. In this manner the process of program evaluation is simplified by allowing main concentration on a few highly significant variables and avoiding the methodological difficulties of interrelating and scaling a host of lesser program characteristics. These lesser characteristics, which often take the form of detailed and specific practices, are acknowledged to be desirable but not necessarily essential. In other words, in the presence of key features that are both necessary and sufficient, the inclusion of lesser characteristics in an evaluation is likely to be cumbersome and wasteful.

Although the concept of key features does not offer a solution to the problem of assigning numerical values of relative importance to varying program features (i.e., weighting), it does guard against the danger of assigning equal merit to characteristics with vastly divergent degrees of worthiness.

The first step in the process of isolating key features consisted of developing a relatively comprehensive list or general features, elements, and processes that represent identifiable dimensions within programs of DEG. This list was based upon those dimensions of differential education which have received considerable and continued emphasis in both the general literature on the gifted and that segment of literature dealing more specifically with programs and program evaluation. The list was submitted to the panel of judges with the requests that (a) they rank, in order of importance, the features which they considered to be the most necessary to a worthy program, and (b) they stop ranking when that number or features had been reached which, if present in an actual program, would assure a program of good quality.

The results of this inquiry were tabulated by means of a pooled frequency rating technique and the features which emerged as being relatively more essential than the others formed the central subdivisions around which the scales of the instrument were developed. The key features, listed in the order that they were ranked by the judges, are as follows: the teacher, the curriculum, student identification and placement, philosophy and objectives, and general staff orientation. These are indicated by the letters A through E on Figure 1.

The fourth step in constructing the instrument was the actual construction of the scales of the experimental evaluative instrument. For each key feature identified by the panel of judges, one or more program requirements was formulated. Program requirements may be thought of as generic expositions of certain theoretical principles or axioms of DEG that are found in the literature and that depict ideally conceived educational practices for exceptionally endowed students. They do not pertain to any given pattern of organization but rather to excellent practices presently operating (either individually or in varying combinations) and practices that can and should be inaugurated in view of the behavioral potential of gifted students. (The focal points of the 11 program requirements are indicated on Figure 1.) It is recognized that in the application of theory to practice, few, if any, existing programs will attain the ideal stature specified in the program requirements. However, the usefulness of the program requirements lies in the function they serve as “benchmarks of quality” or point of departure from which graduated sets of standards could be derived. Viewed in this context, the program requirements themselves may be thought of as principles having particular reference to each program dimension upon which the instrument is focused.

Each program requirement served as a focal idea around which a set of five scale standards was developed. The scale standards are practices or provisions that are derivatives of the respective program requirements and that have been arranged according to positioned degrees of quality along a five interval hierarchy. The verbal tags, *ideal*, *superior*, *commendable*, *neutral*, and *negative* were affixed to each set of scale standards. Numerical values ranging from three to minus one were also assigned. The verbal tags are meant to facilitate constructive interpretation of the instrument. The numerical values similarly affixed also are intended mainly for descriptive or communicative purposes, no attempt having yet been made to construct a numerary value system based on scale theory.

Figure 1

Evaluation Scales for Differential Education for the Gifted (ESDEG)

(Experimental Form, 10-67)

| | | Ideal | Superior | Commendable | Neutral | Negative |
|---------------------|--|-------|----------|-------------|---------|----------|
| 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: GENERAL STAFF ORIENTATION | | | | | |
| Program Requirement | 3: System Wide Support | _____ | _____ | _____ | _____ | _____ |
| KEY FEATURE | C: STUDENT IDENTIFICATION AND PLACEMENT | | | | | |
| Program Requirement | 4: Validity of Conception and Adequacy of Procedures | _____ | _____ | _____ | _____ | _____ |
| Program Requirement | 5: Appropriateness of Relationship Between Capacity and Curriculum | _____ | _____ | _____ | _____ | _____ |
| KEY FEATURE | D: THE CURRICULUM | | | | | |
| Program Requirement | 6: Relevance of Conception | _____ | _____ | _____ | _____ | _____ |
| Program Requirement | 7: Comprehensiveness | _____ | _____ | _____ | _____ | _____ |
| Program Requirement | 8: Articulation | _____ | _____ | _____ | _____ | _____ |
| Program Requirement | 9: Accuracy of Instructional Facilities | _____ | _____ | _____ | _____ | _____ |
| KEY FEATURE | E: THE TEACHER | | | | | |
| Program Requirement | 10: Selection | _____ | _____ | _____ | _____ | _____ |
| Program Requirement | 11: Training | _____ | _____ | _____ | _____ | _____ |
| TOTAL | | X3= | X2= | X1= | X0= | X-1= |
| TOTAL SCORE | | _____ | | | | |

In view of the absence of discrete standard units in the psychological and social sciences, the selection of a five interval ordinal scale represents a compromise between other potentially operable scales that could have been adopted. An effort was made to construct a functional scale that, in addition to differentiating sufficiently between highly desirable practices and those with neutral or negative value for programs of DEG, would recognize and give credit to the sincere efforts of well intentioned people who have attempted to make some provisions for their gifted students. Hence, there are three degrees of positive quality above the merely neutral provision, The negative value is included by virtue of the supportable fact that despite good intentions some practices thought useful for gifted students are actually undesirable or otherwise unjustifiable.

Finally, the experimental instrument was presented to the panel of judges in an attempt to determine whether or not each of the program requirements and scale standards was sufficiently well conceived and structured to allow differentiation in use by qualified observers in the process of evaluating programs of DEG. The judges were asked to accept or reject each item, and in those instances where individual judges believed that improvement in an item was essential to its validity, they were further requested to express their reasoning.

The responses of the judges to this final inquiry were tabulated and a composite of their reactions and suggestions was incorporated into the experimental edition of the instrument (Figure 1).

Preliminary field testing of the experimental instrument has recently been completed by Burnette, Rogge, Plowman, and Kelley, and their reactions which constitute the third and final part of this symposium will hopefully lead to further refinements in the scales.

Conclusion

In conclusion, I would like to emphasize that the instrument here described, or any instrument that is based solely on the judgment and beliefs of individuals, is valid only to the extent that its effectiveness can be demonstrated in actual practice. Thus, the present endeavor is offered as the necessary first step of a continuous movement toward the implementation of practical efforts to evaluate programs for the gifted. The present product must be recognized as a purely experimental instrument until additional investigation and field testing designed to meet the well known mandates of measurement theory are carried out.