

# *Trends in Evaluation*

J. WAYNE WRIGHTSTONE

**How have modern evaluative concepts developed? J. Wayne Wrightstone, director, Bureau of Educational Research, Board of Education of the City of New York, traces their evolution through recent decades.**

THE MODERN CONCEPT of evaluation has evolved gradually in recent decades from a newer philosophy of education and the development of newer techniques of assessing growth and development. The newer philosophy of education has emphasized the responsibility of the educator not only for the development of tangible concepts, information, skills and habits but also the less tangible attitudes, appreciations, interests, powers of thinking and personal-social adaptability of the individual.

As these values or objectives have become clarified and defined in instructional practices and behavior, both formal and informal methods of assessment have been devised to meet the needs. New skills have been stressed by the modern curriculum. These are represented by work-study skills, such as map reading, finding topics in reference books, using an index or table of contents, reading charts, graphs or tables, and knowledge of simple library skills.

## **Modern Evaluative Criteria**

Modern evaluation is characterized by several criteria. First, it attempts to measure a comprehensive range of objectives of the modern school curriculum rather than limited subject-matter achievement only. Second, modern

evaluation uses a variety of techniques of appraisal such as achievement, attitude, personality and character tests. Included also are rating scales, questionnaires, judgment scales of products, interviews, controlled-observation techniques, sociometric techniques and anecdotal records. Third, evaluation includes integrating and interpreting the various indexes of behavior changes so as to construct an inclusive portrait of an individual or an educational situation. For this purpose a comprehensive cumulative record is valuable.

The origins of modern evaluation go back many decades. In the United States the concept of evaluation as related to measurement extends from J. M. Rice about 1900 to the present-day programs. Monroe<sup>1</sup> has distinguished between measurement and evaluation by indicating that the emphasis in measurement is upon single aspects of subject-matter achievement or specific skills and abilities but that the emphasis in evaluation is upon broad personality changes and major objectives of an educational program. From the point of view of its functions, evaluation involves the identification and formulation of a comprehensive range of major objectives of a curriculum,

<sup>1</sup> Monroe, W. S. "Educational Measurements in 1920 and in 1945." *Journal of Educational Research* 38:334-40; January 1945.

their definition in terms of pupil behavior to be realized, and the selection or construction of valid, reliable and practical instruments for appraising major objectives of the curriculum.

### **Standardized Tests Developed**

In 1897 Joseph M. Rice published his studies, entitled, "The Futility of the Spelling Grind." In his study, Rice set up the design of an experiment for comparing the achievement of pupils who spent an inordinate amount of time on spelling drill versus those who spent less time. In order to obtain a measure of achievement he devised a spelling test which represents the first formal type of test in the modern fashion. Rice was interested not only in the measurement of spelling as such, but also in evaluating the strengths and weaknesses of practices and outlooks in the educational program of his time. This represented a break with the previous idea in education, prior to 1900, that learning products were intangible and could be appraised only by the teacher of a particular class.

During this same decade of 1900 we find some initial experimentation with the Binet scale for evaluating intelligence and some beginnings of achievement tests in the basic skills of arithmetic and the language arts. In the decades which follow the initial formal test construction by Rice, one can trace the rise of modern measurement and evaluation.

During the second decade of the century, roughly about 1910, the development of some of the early standardized tests of intelligence and achievement, and the struggle to get them accepted by educators and citizens, were evident.

An important milestone in the use of achievement tests was recognized when standardized tests were employed on a large scale for the first time in the survey of the City of New York in 1911-13. Subsequent to this, standardized tests were used in other large cities. During this period E. L. Thorndike, Leonard A. Ayres, Stuart Courtis, Lewis Terman, Arthur Otis and others were pioneer leaders in the development of tests and measures.

Beginning with the 1920's, tests both of intelligence and achievement were forging ahead rapidly. The use of intelligence tests during World War I for classifying personnel in the armed forces provided an opportunity for the employment of new tests and gave encouragement to the movement.

In 1919 achievement-test batteries and intelligence tests first were issued by commercial publishers for purchase by school systems. Notable among these were such pioneer tests as the Otis Intelligence Test, the Thorndike Scales of Handwriting, various arithmetic tests, and the Stanford Achievement Test batteries. During this period the development of statistical techniques of test analysis also received considerable attention.

### **Evaluation Movement Expanded**

Beginning with 1930, the measurement and evaluation movement expanded and matured. It was during this decade that the Cooperative Test Service began to publish a large number of comparable forms of tests for secondary schools. Many new batteries of tests appeared. During this period also personality tests, including the Rorschach and other projective tech-

niques appeared. Interest inventories were first devised. Attitude scales made their appearance, and sociometric techniques were introduced. Anecdotal records appeared upon the scene. All these efforts represented attempts to find ways and means of assessing not only tangible skills and knowledges, but also less tangible objectives.

It was during this period also that studies evaluating newer practices in elementary and secondary schools were carried on by Tyler, Wrightstone and others. At the college level various studies were conducted by Eurich. At the same time evaluative criteria were set up for accrediting high schools and colleges. These criteria recognized the comprehensive and flexible nature of the curriculum of these schools.

### Techniques Refined

The decade of 1940 has been marked by a maturing and a refinement of the techniques developed during the decade of 1930. Various evaluation studies will indicate the scope and nature of these evaluation programs. In New York City the appraisal of the activity program in the elementary schools<sup>2</sup> involved a comprehensive design of evaluation. In the secondary schools the "Eight-Year Study" of thirty high schools<sup>3</sup> illustrates the broad scope of a modern evaluation program. Currently the Cooperative Study of Evaluation in General Education of the American Council on Education for colleges is

using similar techniques. The Commission on Teacher Education also formulated a broad program of evaluation.

In order to indicate the scope of the objectives measured, Smith<sup>4</sup> reviewed recent evaluation studies and found nine recurring major objectives. These are: mastery of basic skills; ways of thinking; understanding and insight as revealed in social behavior; knowledge gained through conduct on social and personal problems; interests as related to activities; personal initiative; creative power; social and civic attitudes; and post-school vocational competence.

### Evaluation in Eight-Year Study

Most comprehensive in design is the Eight-Year Study directed by Tyler. The report on the evaluation<sup>5</sup> provides a detailed description of the development and application of the evaluation instruments. These instruments were designed to measure the major objectives in the high schools. They are as follows:

1. For aspects of thinking: tests of interpretation of data, application of principles, logical reasoning and nature of proof were constructed.
2. For social sensitivity: tests for application of social values and application of social facts and generalizations to social problems were devised.
3. For civic and social beliefs: scales of social, political and economic beliefs were constructed.
4. For aspects of appreciation in lit-

<sup>2</sup> Morrison, J. Cayce. *The Activity Program: A Curriculum Experiment*. New York: Board of Education of the City of New York, 1941.

<sup>3</sup> Smith, E. R., Tyler, R. W. et al., *Appraising and Recording Student Progress*. New York: Harper and Brothers, 1942.

<sup>4</sup> Smith, Dora V. "Recent Procedures in Evaluation of Programs in English." *Journal of Educational Research* 38: 262-75, December 1944.

<sup>5</sup> Smith, E. R., Tyler, R. W. et al. *Op. cit.*

erature and art: a variety of techniques were employed.

5. For interests: an inventory of personal, social and school interests was devised.

6. For personal and social development: various self-reporting scales and anecdotal records were used.

In addition, various pupil record forms for reading and listening were used. The tests, scales, inventories, questionnaires, check lists, pupil logs and other records were applied as required in each of the thirty schools to gather evidence about the achievement of objectives in this curriculum.

In connection with this study Raths<sup>6</sup> has described and illustrated the newer conceptions of evaluation as applied to evaluating the program of a progressive high school. He showed how various tests and instruments of appraisal for such objectives as interpreting data, applying scientific principles and evaluating data may be used to analyze individual pupil behavior as well as class performance. He illustrated by case studies the integrating of various indexes of behavior into an inclusive appraisal portrait of the individual.

In the elementary school the appraisal of the activity program in the New York City schools<sup>7</sup> measured a variety of objectives such as basic skills in reading, language arts and arithmetic, selected work-study skills such as

map reading, finding references, using an index and interpreting charts, graphs and tables. Interests were evaluated by a specially devised interest inventory. Powers of thinking were measured by special test of interpretation of data and application of generalizations to new situations. In personal and social adjustment various observational techniques, anecdotal records and sociometric techniques were employed. In evaluating creative expression in language a scale for assessing the originality of written compositions was devised.

#### Present-Day Practice

Trends in evaluation in recent decades have evolved from a newer philosophy of education and the development of newer techniques for assessing growth and development. Modern concepts in evaluation require, first, formulation of a comprehensive range of curricular objectives which include not only acquisition of academic information and skills, but also interests, attitudes, appreciations, critical thinking and personal and social adaptability. These objectives should be defined in terms of pupil behavior which indicate growth toward the objective so that appropriate techniques of appraisal may be selected or devised.

Pupil growth and development should be evaluated by both formal and informal tests and techniques. Both types of evaluation are essential to gather evidence of inclusive pupil growth. Newer methods of evaluation will employ tests, rating scales, anecdotal records, interviews, observational techniques, personality tests and sociometric techniques. At some points the pupil can aid in evaluating his own

---

<sup>6</sup> Raths, L. E. "Evaluating the Program of a School." *Educational Research Bulletin* 17:57-84, March 1938.

<sup>7</sup> Morrison, J. Cayce. *Op. cit.* See also: Wrightstone, J. W. "Evaluation of the Experiment with the Activity Program in New York City Elementary Schools." *Journal of Educational Research* 38:252-57, December 1944.

achievement by keeping a record of the books he has read and the radio and television programs he hears and sees.

In order to interpret evaluation data wisely the various items of evidence col-

lected about the pupil should be correlated and integrated into a unified picture of the individual. The relationships among various aspects of pupil growth should be explicitly shown.

---

## *A Region Evaluates Its Elementary Schools*

SARAH LOU HAMMOND and HAROLD D. DRUMMOND

Purposes and techniques of the Southern Association's Cooperative Study in Elementary Education are described by Sarah Lou Hammond, Florida State University, Tallahassee, chairman of the Evaluation Committee, and Harold D. Drummond, George Peabody College for Teachers, Nashville, coordinator of the Study.

"WHY NOT TAKE A NEW LOOK at your school? You may have become so accustomed to the things around you that you cease to see them. Imagine you have never seen your school before and ask yourself, 'Is this a good school for children?'"<sup>1</sup> Last year approximately two hundred elementary school faculties in thirteen southern states were asking themselves questions—searching, penetrating questions designed to help them evaluate the total program of the elementary school. This year probably two or three thousand faculty groups will be so engaged. The Southern Association's Cooperative Study in Elementary Education is responsible for this regional interest in elementary school evaluation.

<sup>1</sup> Unpublished suggestions for revision of *Tentative Edition, Elementary Evaluative Criteria*, developed by Workshops at George Peabody College and A. & I. State College, Nashville, Tenn., Summer 1950.

### Structure of the Study

The structure of the Southern Association's Cooperative Study in Elementary Education is comparatively simple. The Commission on Curricular Problems and Research of the Southern Association of Secondary Schools and Colleges is the organization sponsoring the study. Six other southern agencies are cooperating.<sup>2</sup> In each of the eleven southern states served by the Southern Association a committee has been selected to work for the improvement of elementary education. Committees in Arkansas and Oklahoma are cooperating also, although secondary schools and colleges in these states hold membership in the North Central Association. In each case the committee

<sup>2</sup> See H. Arnold Perry, "The Southern Association's Cooperative Study in Elementary Education," *The High School Journal*, Vol. XXXII, No. 3, May 1949, p. 104.

Copyright © 1950 by the Association for Supervision and Curriculum Development. All rights reserved.