

APPRAISING STUDENTS' REACTIONS
TO AN ENGLISH PROGRAM

Peter Dunn-Rankin
University of Hawaii

Many of us are familiar with curricular programs in which average differences in student achievement between experimental and control groups failed to reach preset levels of significance. We may possibly have realized that the operational definitions of our objectives (generally incorporating standardized objective tests) were inadequate to describe real and important differences between the groups that did occur but were not assumed or immediately reflected in measures of cognitive achievement.

Enthusiasm for short stories, liking for poetry, interest in writing, willingness to read, respect for the teacher, liking for English, etc., are rarely operationally defined. Evaluation of such attitudinal variables can be inferred only in the most superficial way from standardized tests of achievement. Yet the promotion of positive attitudes toward the curriculum is often a primary objective of the instructional program (NCTE Commission on the English Curriculum, 1952, 1956).

Dr. Harold Tanyzer, co-author of the Early to Read materials utilizing the initial teaching alphabet, has remarked that important differences between experimental programs in reading may occur only in the affective domain. That is, a receptive attitude toward reading, writing, and spelling may be the important difference between traditional and experimental programs.

Torsten Husén, who is chairman and technical director of the International Project for the Evaluation of Educational Attainment, measured achievement, in mathematics, of thirteen-year-old students from 12 different countries (Husén, 1966). In addition to the students' responses to a 400-item mathematics test, measures on twenty-six other variables were obtained. These included class size, teachers' salary, and other non-cognitive measures. Husén related that one important non-cognitive-variable accounted for 12% of the variance in performance and was the most effective predictor of success in mathematics achievement. This variable was an index of interest in mathematics.

It is probably not surprising that curriculum motivation is an important variable in achievement. However, the extent of its dominance over other variables in Husén's

study is interesting and lends support to measuring attitudinal variables in order to effectively evaluate achievement in English programs of instruction.

Mill (1960) relates that a child's attitude toward the subject matter often sets up a process of selective attention. He hypothesizes that what a child learns is determined in part by his readiness to receive. Carter (1959) indicates that attitudinal factors are useful predictors of grades and are independent of mental ability tests.

These citations provide a brief rationale to illustrate that attitude measurement is important because the development of satisfaction values in language arts is a primary objective of instruction and because such attitudes may be related to achievement in the cognitive domain.

Most "Project English" curriculum centers have not attempted to appraise students' reactions to their programs. Some centers plan on such measurement in the future. Others feel that attitudinal measurement lies outside the objectives of their experimental curricula. A few have made serious attempts at appraising students' reactions (Steinberg, 1965). One reason evaluators hesitate to measure attitudes is because the subjective judgment involved has a multidimensional character. The student may think the stories in his text are corny but is thankful because the ability level is such that he can read it. When asked how he likes the textbook he combines these reactions to give the text an average rating. In fact, he likes its simplicity but objects to its content, i.e., he has a two dimensional view of the text.

Let me emphasize that there are two kinds of affective domain variables which are particularly relevant to measurement in language arts programs. They can be exemplified by contrasting a person's enthusiasm for reading short stories with his appreciation of short stories as a literary form.

The difference can be related to the Taxonomy of educational objectives in the affective domain (Krathwohl, et al, 1964) in which 3.0 Valuing (behavior is consistent and stable enough to have the characteristics of an attitude) is contrasted with 4.0 Organization (values are organized into a system and interrelationships between the values are determined). My wife, for example, enjoyed reading the "James Bond" books (valuing) but after reading several became critical of their sameness (organization of values).

I make this distinction because it is the concern of this paper to look at students' reactions to English programs at the level of value. Objectives such as

appreciation for short stories and the quality of conceptuality in writing are examples of attitudinal values which are more highly organized (Squire, 1964). Measurement in this area is extremely important but more difficult and beyond the scope of this presentation.

There are many ways in which to assess students' attitudes toward curricular programs. There are, however, probably no better judges as to whether or not various aspects of the curricular program are motivating than the students themselves (Remmers, 1963). What are the ways in which students' reactions to curricular programs can be measured? Most of the effective ways of measuring attitudes involve some sort of scaling technique. These techniques include paired comparisons, rank methods, summated ratings, equal intervals, scalogram analysis, and multidimensional scaling. Other well-known techniques include the semantic differential and Q methodology. Rating scales are used predominantly but such devices as check lists, diaries, inventories, situational appraisals, observational or anecdotal records, and opinion polls are also commonly used.

The technique selected should certainly fit the purpose for which it will be used, but a rating form also needs to be efficient in terms of time and cost. That is, it should be considered in terms of a utility function that relates the value of the information obtained to the time and effort used to obtain it.

As part of the curricular evaluation plan for "Project English" at Florida State University, an attempt was made to evaluate Junior High School students' reactions to three different curricular programs.

Curriculum I is a tri-component approach in which the curriculum is organized into blocks of content with specific units in linguistics, written composition, and literature.

Curriculum II is organized into thematic units which are literature centered and involves humanistic relationships.

Curriculum III is based on a study of cognitive processes involving recognition of form, relation of universals, generalization, and analogies.

Curriculum writers in each order were interested in whether students liked the material they had written.

Administrators wished to compare the receptions that students gave to certain units of each curriculum, such as assignments or literature materials, to see if they differed. Evaluators were interested in validating students' evaluations of different curricula with outside observers' judgments of the effect of such programs.

In an effort to answer these and similar questions the evaluation team at F.S.U. developed an attitude measuring device which may be called a General Semantic Differential. Remmers (1963) has suggested the use of the semantic differential as a widely useful research instrument. Hastings (1964) also felt that an attempt should be made to use the semantic differential in evaluation of the curriculum.

A semantic differential consists of a series of bi-polar adjectives separated by a graphic rating scale. For example, the concept poetry might be rated on a series of scales:

Poetry

good.....bad
valuable.....worthless

heavy.....light
strong.....weak

A factor analysis of students' responses to the scales reveals the number of identifiable and reliably measurable independent dimensions inherent in the scale. We might find, for example, that factor analysis of responses to the four scales listed above reveals two dimensions: Evaluation (items 1 and 2) and Potency (items 3 and 4). The average of a person's scale scores for poetry on the dimension of evaluation can then be used as an index of his attitude toward poetry in that dimension.

Some attempts have been made to use the semantic differential with the concept Teacher (Remmers, 1963). But little use has been made of the instrument in measuring students' attitudes toward other aspects of the curriculum, such as assignments, tests, lectures, class discussions, etc. Della Piana, et al (1965) used the differential to measure reaction to the subject arithmetic and Kerlinger (1965) illustrates its use over the concept school.

The differential created at Project English F.S.U. differed from the typical instrument in two specific ways. The first difference was that the sets of bi-polar adjectives chosen were scored and analyzed over a wide set of

of concepts particularly related to the English curriculum in the Junior High School. These concepts included: assignments, poems, stories, English course, tests, teacher, etc. This was done so that the factor structures of the bi-polar adjectives would remain stable when different concepts were used with the same set of scales. Secondly, the instrument was constructed so that a minimum amount of time was necessary to score and analyze the results.

Four dimensions were identified which were felt to be relevant to the analysis of reactions to curriculum in the language arts. These dimensions and appropriate scales were as follows:

Evaluation:

valuable.....worthless
good.....bad
tasteful.....distasteful
pleasurable.....painful
interesting.....boring

Complexity:

easy.....hard
light.....heavy
simple.....complex

Unusuality:

unusual.....usual
new.....old

Potency:

honest.....dishonest
strong.....weak
complete.....incomplete

Three dimensions, Evaluation and Complexity and Unusuality maintained their factor structure over all the concepts utilized in the pilot studies. The Potency factors were less definite. Perhaps it stretched the imagination to determine what was meant by a dishonest poem. The bi-polar pair "humorous - serious" changed its structure depending upon the concept being rated.

The value of the semantic differential as a comparative instrument can be best illustrated by examining the 3-dimensional representation in Figure 1.

It is not difficult to see that a variety of questions can be asked within a structure of this type.

Specific questions that the Project English staff wanted to answer were:

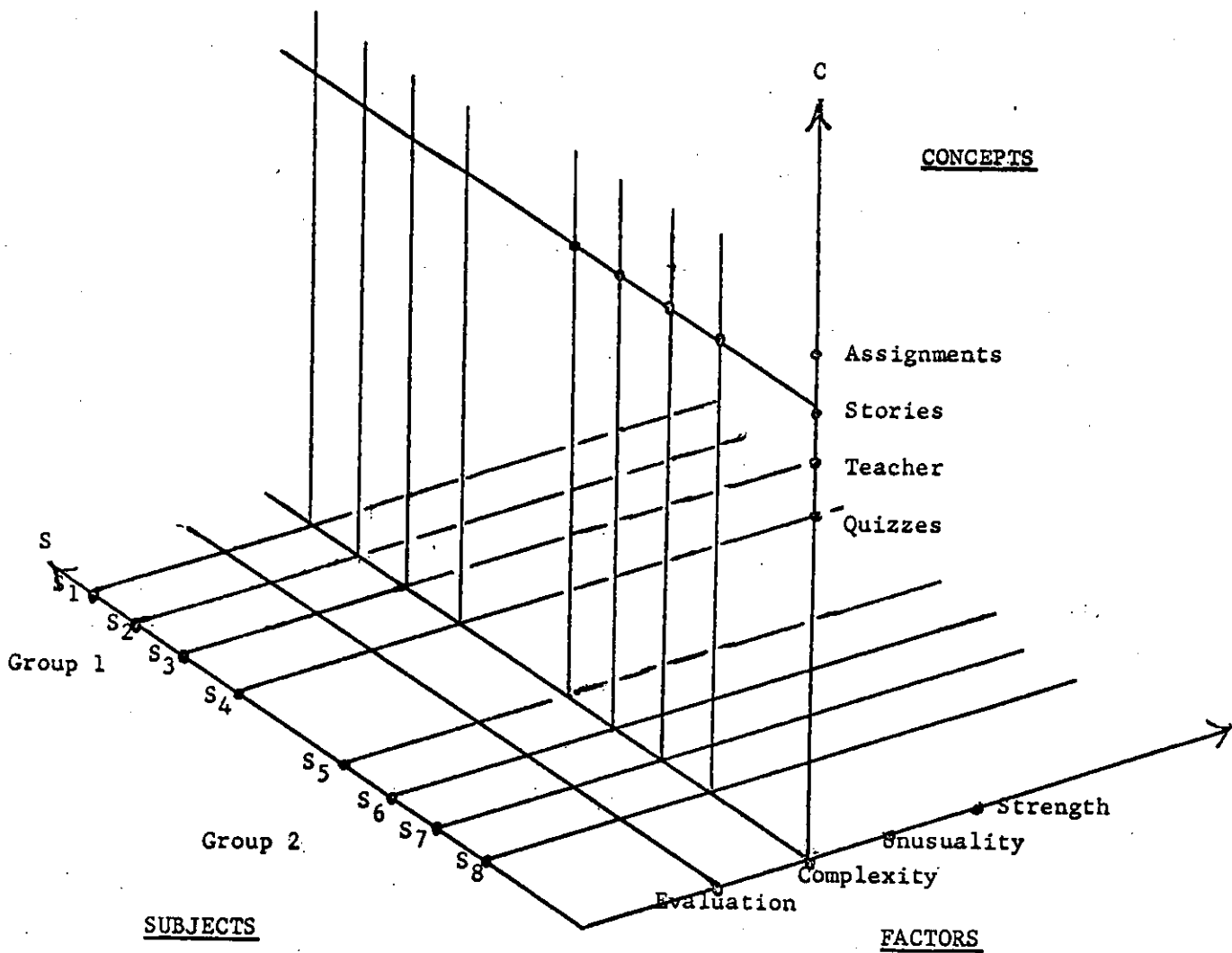


Fig. 1 - A theoretical analysis is illustrated in which two groups' reactions to stories is evaluated on the semantic dimension of simplicity - complexity.

1. Is the attitude of the students toward English programs in each curriculum the same? Are there differences in attitude between the three curricular groups on a single concept or on a single factor or both?
2. Did the students in one curriculum see their program as being more valuable, more unusual, or simpler than students in other programs?
3. Did students indicate any differences between the dimensions of factors of attitude? Were there, for example, programs seen as highly unusual, complex, but not very valuable?
4. Did the students differentiate some aspects of the curriculum as more valuable, complex, or unusual than other areas of the curriculum? Were, for example, written assignments more complex and difficult than reading assignments? Which were viewed as more valuable?

In May, 1964, 1200 students in six different Junior High Schools responded to the instrument. After administration the answer sheets were scored on the IBM 1230/534 optical scorer. At the same time, cards were punched with the students' responses. This deck of cards was processed through a Fortran routine and new cards with average factor scores produced.

Table 1 illustrates the results of multiple comparisons between the three curricula. In this case the possible scores run from 0-4 and the factors are in the order: valuable-not valuable; simple-complex; unusual-usual. The smaller the mean the more favorable was the response.

The Newman-Kuels sequential range test (Ryan, 1959) was used to test the significance of the difference between any two curriculum means. The values in the table indicate the differences between the means of the row and column curricula. If the difference is starred it indicates that a difference as large or larger than the one tabled would only occur by chance five times in one hundred.

The results of these comparisons are as follows:

1. On the factor of Evaluation CII, the thematic units curriculum, is seen as significantly more valuable than CI, the tri-component curriculum.

Table 1

Multiple Comparisons Between Three Curricula over the
Concept English Course. Dimensions of Evaluation,
 Complexity and Unusuality are shown.

Evaluation			CI	CIII	CII
	Means		1.1125	.9966	.8983
Tri-component	CI	1.1125	-		
Cognitive process	CIII	.9966	.1159	-	
Thematic units	CII	.8983	.2142*	.983	-
Complexity			CII	CI	CIII
	Means		2.0442	1.9529	1.7899
Thematic units	CII	2.0442	-		
Tri-component	CI	1.9529	.0913	-	
Cognitive process	CIII	1.7899	.2543*	.1630*	-
Unusuality			CIII	CI	CII
	Means		1.4776	1.1933	1.1271
Cognitive process	CIII	1.4776	-		
Tri-component	CI	1.1933	.2843*	-	
Thematic units	CII	1.1271	.3505*	.0662	-

*Significant at .05 level

2. On the factor of Complexity CIII, the cognitive processes curriculum, is seen as less complex than either CI, the tri-component curriculum, or CII, the thematic units.
3. On the factor of Unusuality CI, the tri-component curriculum, and CII, the thematic units, are seen as more unusual than CIII, the cognitive processes.

When achievement measurements for these individuals are also taken, the meaning of these differences may become more apparent and more valuable.

It would appear that the general semantic differential can be effectively used as an instrument for the evaluation of students' reactions to English programs for the following reasons:

1. It is flexible. The questions that can be asked of the data are almost too numerous to relate.
2. It is simple to administer. Approximate administration time is 10-15 minutes.

3. It can be scored and analyzed quickly and accurately.
4. It is an acceptable measuring device. The subtlety of most of the bi-polar adjectives may make it less threatening than other rating scales.

References

- Carter, Harold D. Improving the prediction of school achievement by use of the California Study Methods Survey. Educational Administration and Supervision, 45, September 1959.
- Della-Piana, Gabriel M. (Ed.) Sequence characteristics of text materials and transfer of learning: Part II Appendices. Salt Lake City: University of Utah, 1965.
- Hastings, J. Thomas. Innovations in evaluation for innovation in curriculum. In Steinberg, Edwin R. et al (Ed.), Curriculum development and evaluation in English and social studies, Cooperative Research Project, No. F-041. Pittsburg: Carnegie Institute of Technology, 1964.
- Husén, Torsten. International project for the evaluation of educational achievement, Part I: A comparative study of outcomes of mathematics instruction in twelve countries. New York: John Wiley and Sons, in press, 1966.
- Kerlinger, Fred N. Foundations of behavioral research. New York: Holt, Rinehart and Winston, Inc., 1965.
- Krathwohl, David R.; Bloom, Benjamin S.; and Masia, Bertram B. Taxonomy of educational objectives, Handbook II: Affective domain, David McKay Co., 1964, pp. 176-185.
- Mill, Cyril R. Attitudes affect pupils' learning. Educational Leadership, January 1960, p. 212.
- National Council of Teachers of English, Commission on the English Curriculum. The English language arts. New York: Appleton-Century-Crofts, 1952.
- National Council of Teachers of English, Commission on the English Curriculum. The English language arts in the secondary school. New York: Appleton-Century-Crofts, 1956.

- Remmers, H. H. Introduction to opinion and attitude measurement. New York: Harper and Brothers, 1954, p. 6.
- Remmers, H. H. Rating methods in research on teaching. In Gage, N. L. (Ed.) Handbook of research on teaching. Chicago: Rand McNally, 1963.
- Ryan, T. A. Multiple comparisons in psychological research, Psychological Bulletin, 1959, 56, 26-47.
- Squire, James R. The responses of adolescents while reading four short stories. NCTE Research Report No. 2. Champaign: National Council of Teachers of English, 1964.
- Steinberg, Erwin R. et al. An evaluative study of a senior high school curriculum in English for able college bound students: Vol. IV Evaluation. Pittsburg: Carnegie Institute of Technology, 1965.