This study tested the hypothesis that teacher involvement in curriculum evaluation facilitates curriculum implementation. Curriculum implementation is defined as a function of the elements(s) of a particular referent curriculum. A random sample of 24 teachers, eight taken from each of three grade levels, served as subjects(S's) for experimental and control groups. S's evaluated a social studies curriculum through a series of meetings with the director of Instructional Services. Observations were made over a 10-week posttest period, and teachers were rated (using a modified version of the OSCAR 4V) on teacher-pupil verbalizing relevant to the curriculum. Findings indicate that curriculum implementation was facilitated differentially by degree of involvement in curriculum engineering. (JB)
AN EXPERIMENTAL STUDY OF THE EFFECTS OF TEACHER PARTICIPATION IN A CURRICULUM ENGINEERING TASK ON A DIMENSION OF CURRICULUM IMPLEMENTATION

James R. Gress
The Ohio State University

Introduction

Considerable time, effort, and financial resources have been allocated and spent in the decade just past in efforts to develop new curriculums. This is reflected in the efforts of both the so-called national curriculum projects and many groups in local school districts. However, the expenditures invested have not resulted in systematically implemented curriculums and the changed emphases in objectives, subject-matter, and instructional strategies presumably intended.

Alexander (1962) stated more than a decade ago that more is known about the extent to which curriculum proposals are made than the extent of their implementation. Indeed, one recent study (Schwartz, 1971) is typical of the emphasis -- in research as well as in practice -- on origination and dissemination of new curriculums rather than upon their implementation and integration into instructional systems.

Some of the variables related to curriculum implementation have been...
identified and categorized (Gress, 1972, pp. 28-33). Those variables include:

1. characteristics of the curriculum itself, 2. characteristics of the teacher as a unit of implementation, and 3. characteristics of other instructional inputs (i.e., the resources and constraints of a given instructional situation).

And the teacher's relationship to a curriculum system, as a variable associated with curriculum implementation, has provided the basis for some research. However, little experimental evidence exists of the effects upon curriculum implementation of particular teacher relationships to curriculum systems as defined by given curriculum engineering tasks.

The Problem

Teacher involvement in such engineering tasks as curriculum planning and evaluation constitute important determinants of curriculum implementation and consequent instructional change (Anderson, 1965, pp. 56-61; Beauchamp, 1964, pp. 355-356; Saylor and Alexander, 1966, pp. 405-437). Perhaps the earliest report to that effect was among some conclusions offered by Saylor in 1941 (pp. 212-221). In his study of a curriculum system, Saylor concluded that there was general curriculum adoption and implementation where teachers participated in curriculum planning and in the construction of instructional materials.

Heusner (1964) and Krey (1969) both concluded that curriculum implementation is dependent, in part, upon the teacher's recognition of the need for change and consequent implementation. In addition, McQuigg (1963) and Edwards (1969) have supplied research evidence to support their common assertion that the teacher's understanding of a curriculum, its relevance to an instructional situation, and his role in the implementation process is a determinant of change. Langenbach (1970) concluded that teachers who had participated in curriculum planning had more positive attitudes toward curriculum use than teachers who had never participated. Johansen (1965) offered data to support his thesis that teacher participation in curriculum planning increased the probability of curriculum
implementation if the teacher perceived that he was influential in curriculum decision-making. Black (1968) found that teacher participation in curriculum evaluation was significantly related to differences in curriculum implementation. Teague also asserted that "...teachers will spend more time and effort in programs they helped plan, participated in, or had an opportunity to evaluate" (Teague, 1963, p. 4).

This study was undertaken to gather experimental evidence to support a thesis relating teacher involvement in decision-making with curriculum implementation. The study posed the hypothesis: Teacher involvement in curriculum evaluation facilitates curriculum implementation.

Research Design

For this study, a curriculum was conceptualized, and curriculum implementation was operationalized, within the systems of schooling framework provided by Beauchamp (1968, pp. 110-144), the definition of curriculum implementation being a function of the element(s) of a particular referent curriculum. The design features of a curriculum might include an outline of subject-matter, a statement of goals and/or specific objectives, a rationale for the curriculum itself, and an evaluation schema (Beauchamp, 1968, pp. 83-85). In addition, many curriculums contain descriptions of instructional activities, bibliographies of instructional materials, and descriptions of classroom instructional strategies (Beauchamp, 1963).

The New Trier Township K-8 Social Studies Curriculum, used as a vehicle for this study, included many of the elements identified here. Since curriculum implementation could be operationalized in terms of the extent to which empirical data might be gathered in an instructional situation based on the element(s) of the curriculum's design, classroom verbal interaction was defined in this study in terms of a description of the instructional strategies contained in The New Trier Township K-8 Social Studies Curriculum (Chapter 3). Figure 1 displays the relationship between characteristics of instructional discourse described in this
social studies curriculum and twelve scales of teacher-pupil classroom verbal behavior derived by combining selected categories of the OScAR 4V (Simon and Boyer, 1970, pp. 13:1-4). (Interpreting curriculum implementation as a function of verbal behavior in the classroom has also been done by Hawthorne (1968; 1971) and Hart (1972)).

Characteristics of Instructional Discourse Described in the Social Studies Curriculum

The basic mode of discourse is questioning.

The teacher rarely "tells" his students what they need to know.

The teacher's questioning is not directed at a single student response.

The teacher encourages active student participation in instruction.

The teacher shows empathy in his reinforcement of students' ideas, statements, and questions.

Scales of Verbal Behavior Derived from the OScAR Categories

+ Teacher Substantive Interchanges (TS)
- Teacher Statements (TS)
+ Teacher Problem-Structuring (PRB)
- Teacher Lecturing (LCT)
+ Divergent Questioning (DVG)
- Convergent Questioning (CVG)
+ Pupil Statements (PS)
+ Pupil-Initiated Behavior (PUP)
- Teacher-Initiated Behavior (TCH)
+ Teachers' Responses (WEX)
- Teacher Considering Behavior (CSD)
- Teacher Rebuking Behavior (RBK)

Figure 1
Characteristics of Social Studies Instruction and Related Verbal Behavior Scales

For the study, a stratified random sample of twenty-four teacher-class combinations was selected from teacher-class units in Elementary School District 39 (Cook County, Illinois), eight teacher-class units at each of three grade levels assigned randomly to experimental and control groups. The district's Director of Instructional Services, the administrator responsible for the district's curriculum
system, also participated in the study.

The treatment for teachers in the experimental group consisted of participation in an evaluation of the social studies curriculum used. The evaluation was accomplished through a series of formal and informal meetings involving the district's Director of Instructional Services with the teachers. These experiences were structured according to selected principles of organization theory (Bennis, 1966; Benne, 1949, pp. 204-207; Jensen, 1969, pp. 103-208). The experimental treatment was initiated in December, 1970, and culminated in April, 1971, with the teachers' formulation of the evaluation. The treatment for teachers in the control group consisted of no participation in the evaluation tasks.

A modified version of the OSCAR 4V (Medley and Mitzel, 1963, pp. 278-286; Medley and Hill, 1968) was employed to assess the dimensions of teacher-pupil classroom verbal behavior relevant to the curriculum. Each teacher who participated in the study was observed during social studies instruction on four different occasions. The observations were made during a period of ten weeks after the experimental treatment had been initiated. Each observation lasted for one-half hour which included time for observer orientation to the instructional setting.

For each teacher, scores were computed for each of the twelve OSCAR scales described. The reliability for each of the twelve scales was estimated by the three-way analysis of variance technique described by Medley and Mitzel (1958, pp. 23-25). Teachers, observations, and OSCAR categories were considered fixed effects. Table 1 summarizes the ANOVA findings and the reliability indices derived.

To test for differences in teachers' total scores on the twelve scales of verbal behavior, use of a multivariate analysis of variance (MANOVA) was planned (Bock and Haggard, 1968, pp. 100-142; Rulon and Brooks, 1968, pp. 72-76). Tests of the assumptions of the MANOVA model were made. Multinormality of population variables (Morrison, 1967, pp. 125-168) was tested indirectly by probit analyses.

The author wishes to acknowledge the assistance of John McConnell, presently a Northwestern University doctoral candidate, in the analysis of data.
TABLE 1

Reliability Co-efficients Derived By Analysis Of Variance for Twelve Scales of Verbal Behavior

<table>
<thead>
<tr>
<th>Scale</th>
<th>Hoyt R</th>
<th>Scale</th>
<th>Hoyt R</th>
<th>Scale</th>
<th>Hoyt R</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCH</td>
<td>0.597</td>
<td>CVG</td>
<td>0.703</td>
<td>PS</td>
<td>0.716</td>
</tr>
<tr>
<td>PUP</td>
<td>0.475</td>
<td>DVG</td>
<td>0.739</td>
<td>WEX</td>
<td>0.705</td>
</tr>
<tr>
<td>LCT</td>
<td>0.756</td>
<td>TS</td>
<td>0.527</td>
<td>CSD</td>
<td>0.450</td>
</tr>
<tr>
<td>PRB</td>
<td>0.747</td>
<td>TSI</td>
<td>0.772</td>
<td>RBK</td>
<td>0.823</td>
</tr>
</tbody>
</table>

a, p < .05 Among Teachers
b, p < .05 Among Categories
c, p < .05 Among Observations

(Goulden, 1952, pp. 394-417) of the individual distributions of scores. The distributions of scores on the PRB, PS, CSD, and RBK scales were found to depart significantly from normality. Homogeneity of dispersion (Hope, 1968, pp. 29-30; Bock, 1966, pp. 810-840) was tested indirectly by applying the $F_{max}$ criterion (Winer, 1962, pp. 92-95) to the within-groups variances for each of the twelve scales. Means and standard deviations of teachers' scores on the twelve scales are displayed by treatment-group and grade-level in Table 2. Within-groups variances of teachers' scores produced by the PRB, PS and WEX scales were found to be nonhomogeneous. On the basis of these findings and by examining the intercorrelations of scores on the twelve scales for the experimental and control groups, five of the scales of verbal behavior were excluded from further analysis. Data for the TCH, PUP, TS, TSI, CVG, DVG and LCT scales were used in the main analysis. To characterize significant multivariate differences, use of multiple linear discriminant functions was planned (Rao, 1952, pp. 237 248; Bock and Haggard, 1968, p. 117; Morrison and Art, 1967).

Additional statistics derived from various analyses of data for each of the twelve verbal behavior scales is available from the author upon request.
### TABLE 2
Means And Standard Deviations Of Teachers' Scores
On Twelve Scales of Teacher-Pupil Verbal Behavior
By Treatment Group And Grade Level

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>Grade Level</th>
<th>N</th>
<th>TECH</th>
<th>PUP</th>
<th>LCT</th>
<th>PRB</th>
<th>TS</th>
<th>TS1</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>2</td>
<td>4</td>
<td>M</td>
<td>84.75</td>
<td>25.75</td>
<td>24.50</td>
<td>3.25</td>
<td>56.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>17.17</td>
<td>7.93</td>
<td>8.34</td>
<td>3.20</td>
<td>15.06</td>
</tr>
<tr>
<td>E</td>
<td>5</td>
<td>4</td>
<td>M</td>
<td>78.00</td>
<td>29.25</td>
<td>14.00</td>
<td>5.00</td>
<td>43.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>5.71</td>
<td>10.43</td>
<td>6.48</td>
<td>5.09</td>
<td>14.08</td>
</tr>
<tr>
<td>E</td>
<td>8</td>
<td>4</td>
<td>M</td>
<td>85.50</td>
<td>21.25</td>
<td>31.00</td>
<td>5.00</td>
<td>65.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>5.91</td>
<td>11.23</td>
<td>13.34</td>
<td>2.58</td>
<td>5.31</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>4</td>
<td>M</td>
<td>93.50</td>
<td>24.00</td>
<td>11.25</td>
<td>3.75</td>
<td>46.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>20.02</td>
<td>7.30</td>
<td>9.21</td>
<td>2.36</td>
<td>14.61</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>4</td>
<td>M</td>
<td>83.75</td>
<td>38.00</td>
<td>20.75</td>
<td>3.00</td>
<td>56.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>13.15</td>
<td>24.56</td>
<td>18.71</td>
<td>3.55</td>
<td>17.25</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>4</td>
<td>M</td>
<td>45.25</td>
<td>55.25</td>
<td>15.75</td>
<td>1.75</td>
<td>35.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>18.08</td>
<td>17.23</td>
<td>10.07</td>
<td>.50</td>
<td>9.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CVG</th>
<th>DVG</th>
<th>PS</th>
<th>CSD</th>
<th>RBK</th>
<th>WEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>2</td>
<td>4 M</td>
<td>15.75</td>
<td>12.75</td>
<td>7.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>8.26</td>
<td>7.76</td>
</tr>
<tr>
<td>E</td>
<td>5</td>
<td>4 M</td>
<td>18.00</td>
<td>14.00</td>
<td>11.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>4.24</td>
<td>10.13</td>
</tr>
<tr>
<td>E</td>
<td>8</td>
<td>4 M</td>
<td>13.00</td>
<td>5.75</td>
<td>7.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>7.61</td>
<td>2.75</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>4 M</td>
<td>25.75</td>
<td>19.25</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>9.21</td>
<td>5.56</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>4 M</td>
<td>19.25</td>
<td>6.75</td>
<td>14.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>7.41</td>
<td>5.37</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>4 M</td>
<td>5.50</td>
<td>4.25</td>
<td>40.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>7.55</td>
<td>7.22</td>
</tr>
</tbody>
</table>
Results

Based on tests of the assumptions of the MANOVA model, data pertaining to seven scales of teacher-pupil classroom verbal behavior were subjected to a $2 \times 3$ multivariate analysis of variance. The results of the analysis are summarized in Table 3.

TABLE 3
Multivariate Analysis of Variance Of Teachers' Scores On Seven Scales of Teacher-Pupil Verbal Behavior

<table>
<thead>
<tr>
<th>Source</th>
<th>Tests of Roots Using Wilks' Lambda</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$df_{HYP}$</td>
<td>$df_{ERR}$</td>
</tr>
<tr>
<td>Treatment Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roots 1 - 1</td>
<td>7.0</td>
<td>12.0</td>
</tr>
<tr>
<td>Grade Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roots 1 - 2</td>
<td>14.0</td>
<td>24.0</td>
</tr>
<tr>
<td>Roots 2 - 2</td>
<td>6.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roots 1 - 1</td>
<td>14.0</td>
<td>24.0</td>
</tr>
<tr>
<td>Roots 2 - 2</td>
<td>6.0</td>
<td>12.5</td>
</tr>
</tbody>
</table>

The result of the F-test of treatment group differences not significant ($p > .05$), but the results of the F-tests of differences among grade levels and differences attributable to treatment group-grade level interaction effects were significant ($p < .05$). It was concluded that there were no differences between teachers attributable only to treatment effects, but that there were grade level and interaction differences.

The data were subjected to post hoc discriminant analyses for grade level and interaction effects. Table 4 summarizes the results of the analysis for grade level effects. Since the stepwise procedure utilized selected only one variable, TS1, for the resultant linear function, no further discriminant analysis.
was possible. Further analysis using Scheffe's comparisons (Ferguson, 1966, pp. 295-297) revealed that differences among teachers in teacher-initiated substantive interchanges (TSI) could be explained in terms of differences between second and fifth versus eighth grade teachers, and it was concluded after further examination of the data that eighth grade teachers tended to use questioning as a basic mode of instructional discourse to a significantly lesser extent than did the other teachers.

Table 5 summarizes the results of the discriminant analysis for interaction effects. Two discriminant functions were generated. Each of those functions contained co-efficients for scores in the TS and TSI scales. Figure 2 displays the discrimination of the group centroids generated by the functions. The first function discriminated between eighth grade and other teachers. The second discriminated further between treatment groups. Re-examination of the data revealed that teachers in the experimental group at each grade level scored consistently higher on the TSI scale than did their counterparts in the control group.
Since Questioning as a basic instructional mode of discourse was implicit in the curriculum design, it was concluded that second and fifth grade teachers implemented the curriculum to a greater extent than did eighth grade teachers. But, more importantly here, it was also concluded that the experimental group tended to utilize that mode of discourse to a greater extent than did the control group. And, since fifth grade teachers also utilized an antithetical mode of discourse (TS) to a significantly lesser extent, it was concluded that the experimental treatment was differentially effective among teachers at different grade levels.
Figure 2

Two-Dimensional Discrimination of Six-Group Centroids for Multivariate Verbal Behavior

E = Experimental Treatment Group
C = Control Treatment Group

2 = 2nd Grade Teachers
5 = 5th Grade Teachers
8 = 8th Grade Teachers
Discussion

It was concluded that curriculum implementation, as defined, was facilitated to some extent, if differentially, for teachers involved in the curriculum engineering task described. However, the study suggests that generalizing from the limited results of the study to the effects upon curriculum implementation potentially attributable to on-going teacher involvement in curriculum engineering decision-making is necessarily constrained by:

1. The nature and extent of the curriculum engineering task(s);
2. The type and number of teachers involved in the decision-making process;
3. The nature of the curriculum and its design element(s).

First, teachers, in this study were involved in a specific decision-making task -- evaluating a pilot curriculum -- for the purpose of making a decision about its adoption and/or revision. In defining participation, other decision-making tasks might also be operationalized, e.g., initial design of a curriculum, on-going evaluation of an already adopted curriculum, negotiating for the implementation of specific instructional objectives. Perhaps, as is suggested by the results of the study, different tasks are differentially effective with different teachers. Perhaps involvement in a particular combination of tasks optimizes implementation. Further, the question of extent, both over time and in degree, of involvement is a factor to be considered. In this study, involvement in decision-making was uniformly and arbitrarily defined and the period of involvement limited to several months. Perhaps other types of involvement in decision-making would have yielded different results. Perhaps, given the nature of the curriculum, the length of time during which teachers were involved was not adequate for demonstrable change or perhaps the lapse of time between involvement and potentially demonstrable change precluded the gathering of stronger experimental evidence in this study.

Secondly, in this study, twenty-four teachers from a school district in an upper-middle class socio-economic community were involved. Perhaps variables relate to unique characteristics of that group, or to their school district and its other policies, or to the population of students represented contributed to the particular effects of involvement in decision-making identified. Or, given a larger sample,
more demonstrable effects might have been shown. Or, perhaps involving all teachers
in a school district in curriculum decision-making, if indeed this is a viable
principle of curriculum engineering, may generate more conclusive results. Or, as
is suggested perhaps grade level effects are an important variable over all population
'scale', the nature of the curriculum itself seems important here. Faray (1970)
has reported that the nature of the subject-matter is an important variable in the
implementation of a curriculum, and it seems relevant that the social studies curri-
culum used as a vehicle for this study dealt with many controversial issues (as it
may well should have) including race relations, environmental concerns, and the like.
And perhaps the particular combination of design features of a curriculum facilitate
per se its implementation. Perhaps, while teachers may not have utilized recommended
instructional strategies, they did cover the subject-matter intended and to that
extent may be said to have implemented the curriculum; and perhaps it is along that
dimension that the effects of involvement in decision-making would have been demon-
strated.

Since it seems reasonable to conclude that curriculum implementation is a
function of a number of interrelated variables, the need for additional study in
this area based on some conceptual framework and utilizing a syntactical structure
based on experimental evidence seems warranted. This is necessary for two reasons,
at least. The evolution of descriptive curriculum theory is not possible without
taking this approach to a study of the variables involved. And, rational curriculum
engineering in schools is dependent upon replacing the prescriptions of the past with
valid descriptive theory.
References


Benne, Kenneth D. Democratic ethics in social engineering. Progressive Education, 1949, 26, 204-207.


