DOCUMENT RESUME

ED 099 304 SO 008 013

AUTHOR Rosenshine, Barak
TITLE Curriculum and Other Contextual Variables.
PUB DATE 74
NOTE 25p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS Academic Achievement; *Content Analysis; Course Content; Curriculum; Curriculum Development; *Curriculum Evaluation; *Curriculum Research; Educational Research; Research Criteria; Research Methodology; *Research Needs; *Research Opportunities; Social Studies; Teaching Procedures

ABSTRACT

This document examines contextual variables and how they affect instructional processes and educational outcomes. One contextual variable which has not received enough attention is the curriculum-material package. Curriculum packages represent potential experimental treatments. Possible research questions are whether these materials and instructions are suitable, whether they are used properly, and whether the outcomes are the ones which are expected. Once the effectiveness of a curriculum package is analyzed, the next research step is to assess the content and how it affects curriculum results. Research to date indicates that different curriculums produce different patterns of achievement and that these patterns are strongly influenced by the content emphasized in the curriculum. Further research into this area might study the influence of content, time available for instruction, student attention to task, opportunity to learn, and teacher emphasis on student achievement. (Author/DE)

September 1, 1974

CURRICULUM AND OTHER CONTEXTUAL VARIABLES[1].

Barak Rosenshine
University of Illinois

In this paper a number of topics on the possible importance of a curriculum package are explored. The first topic is a discussion of contextual variables, followed by an argument for the importance of a curriculum package in aiding instruction. Next is a discussion of how knowledge of curriculum could be used in process outcome research and a discussion of previously neglected variables such as the amount of help which a curriculum might provide for a teacher. The paper concludes with a discussion of two related topics: the importance of considering content covered and the importance of doing research to eliminate unproductive research ideas.

The Number of Contextual Variables

The argument for considering contextual variables is that the correlation between instructional processes and educational outcomes will be different when one is stratified by the proposed contextual variable. For example, it is frequently suggested that if classrooms were stratified by the pupils' socioeconomic status the correlations between instructional processes and student outcomes would be different

---

1. Paper delivered to the annual meeting of the AERA: Chicago, 1974

within each SES level. Similarly, it is suggested that correlations will differ across grade levels, across subject areas, according to the outcome measure being used, according to the location of the school (e.g. urban, rural, suburban), according to the SES status of the pupils, and according to any number of suggested contextual variables.

The potential importance of context can operate in a number of ways. Context can be a stratifying variable in that certain process-outcome correlational relationships may be stronger or weaker in different contextual combinations. It can also operate as a facilitative variable in an experiment. In such cases, one studies whether certain acquired teaching skills (or certain curriculum programs) operate more effectively in one context or another. The effects of training teachers in asking higher level questions, for example, may be stronger in some contexts than in others, or, other teaching procedures such as highly structured textbooks or pupil-directed inquiry programs may be differentially effective in different contexts.

At the same time that one identifies contextual variables and hopes that research in this area can be used to provide clearer statements on the generality of teaching skills or instructional programs, one notes that such promise is accompanied by problems. Namely, the number of possible contextual variables is so large that it can become extremely difficult to disprove the latest educational gimmick because the innovators can always claim that if an innovation has not been shown to be effective it is because it was tried out in an inappropriate contextual setting.

The number of possible contextual variables is bewildering. Consider the following list of possible contextual variables (and the number of possible divisions within each variable):

Pupil variables

socioeconomic status
learning style
level of cognitive development
sex
ethnic background
linguistic background
age
attitude toward school

Teacher variables

beliefs, perceptions, and personality
stability of classroom interactions
stability in obtaining similar pupil outcomes across years
experience
socioeconomic status
sex
ethnic background
linguistic backgroumd
age

Curriculum variables

grade level
subject area
type of curriculum (e.g. controlled inductive, free inductive)
outcome measure(s)

Community variables

location
SES
ethnicity

Organizational variables

school size
classroom size
school organization
classroom organization
resources available (physical, financial, staff)

The number of combinations of contextual variables is staggering. If one divided each of the above variables into the appropriate divisions (e.g. two levels of sex; five types of subject area) there would be a total of well over two million possible cells. If one attended only to a subset of these variables such as subject area, grade level, school location and SES there would be 270 cells. But the total of all studies relating classroom behaviors to student outcomes is less than 100, and this number grows at the rate of only eight to twelve studies per year.

Discussion of contextual variables suggests that we be cautious in our generalizations and limit them to those contexts in which they were tested. Yet, the contexts which have been suggested and which will be suggested by others encompass far more situations than studies.

This problem of contextual variables can lead to thinking in opposite directions. On one hand, one hopes that most contextual variables can be safely disregarded and that the generalizations which are developed will hold across most situations. Alternatively, the generalizations about instructional processes and student outcomes which have been developed to date are not particularly strong, leading to the converse hope that if some contextual variables were attended to the results would be stronger. The effect of raising these contextual variables is to force a researcher or reviewer to state the limits of the generalizations he wishes to make. Thus, if one wished to draw generalizations from the research on a teacher's use of structuring statements, for example, it would seem reasonable, at this time, to state the contexts to which such a generalization might be applied and the contexts in which such a generalization has not been tested.

As examples of the use of contextual variables in research, consider some results on two contextual variables: social class and teacher stability in producing pupil achievement. In the first case the results support using the contextual variable; in the second case they don't.

Social class as a contextual variable. The research by Brophy and Evertson (1973, 1974) illustrates how one might look at contextual variables. In their studies of instructional behaviors and student

achievement, they divided their teachers into two groups--those who taught in lower class schools and those who taught in middle class schools. In their analyses of both the first year and the second year data there were a large number of variables for which the magnitude and direction of the correlations were different in the two contexts, a sufficiently large number of variables to suggest that classification by SES context is worthwhile--at least for reading and mathematics gain in the second grade.

Teacher stability as a contextual variable. Brophy and Evertson also believed that if they limited their sample to that of stable teachers--those teachers who were consistent in obtaining the same amount of student gain across several years--their process-outcome correlations would be higher than if they had merely selected a random sample of teachers. This hypothesis was not confirmed--at least for reading and mathematics gain in the second grade.

Other contextual variables. There is a need for a closer look at the actual importance of many of the contextual variables which are presented as potentially important. I did an informal look at two contextual variables--grade level and subject area--by consulting the tables in my book on teaching behavior and student achievement (1971). Each study in those tables is listed by grade level and subject area, and one can use this information to reorganize the tables to determine whether the overall results would be different or distinct for any grade level or subject area. When my students and I did this, we found, first, that there were a large number of blanks for most of the combinations of grade level and subject areas. We also found that none of the results which were obtained were particular to any combination of grade level or subject area. Thus, the overall results on teacher questions, for example, are unchanged if one divides the tables on questions according to grade level or subject area; the results on clarity or on task orientation appear significant across all grade levels and subject areas reported.

Curriculum Packages as a Contextual Variable

After raising the horror of a two million cell contextual variables matrix, I am bewildered by my audacity when I proceed, as I shall, to suggest that there is an important contextual variable which has not received enough attention: curriculum-materials packages. Let me sketch, hesitantly, the argument for this variable.

First, consider the possibility that a teacher does not teach fourth grade reading, for example. Rather, she is teaching fourth grade MacMillan readers, or fourth grade SRA, or Ginn, or Open Court. The instruction is embedded within the context of instructional materials.

Consider also the possibility that a curriculum program can serve as a tool to enable a teacher to accomplish more than he could without the materials. One example of such a curriculum materials package (or an instructional package) was developed about 1910 by Montessori (Evans, 1971) and is a superb example of this invention. The Montessori Method included specific, self-correctional materials (e.g. the brown stair, the pink tower, the golden beads), specific instructions for teacher interactions with the child (e.g. vocabulary development had a three-period sequence: naming, recognition, and pronunciation), and general instructions for teachers (e.g. collaborative work with the child, avoidance of "don'ts," emotional support). The instructional materials, sequencing, provision for corrective feedback, and specific and general instructions to teachers contained in the Montessori materials are quite different from the usual practice of providing a teacher with only a set of books, a syllabus, and vague objectives. The major advantage of the Montessori package, or any curriculum materials package, is that it may enable a teacher to accomplish ends which would not be accomplished without these materials.

The advantage claimed for this approach is that curriculum-materials packages represent potential experimental treatments and also provide a teacher with means to accomplish more than he could without the materials. Whether these materials and instructions are suitable, whether they are used properly, and whether the outcomes are the ones which are expected,

are the research questions. Although there has been much development of curriculum materials packages, there has been relatively little study of the research questions.

The Sullivan programmed reading materials, which are used in kindergarten and first grade as part of two Planned Variation Follow Through (PVFT) programs--The Pittsburgh model and the Kansas model--are additional examples of facilitative materials which provide help for a teacher. Among the activities handled in this program are: locating materials, sequencing instruction, and providing feedback to children. The responsibility for developing the instruction and testing the efficacy of the approach lies with the developer who has, presumably, solved many of these instructional problems in previous trials.

The results of the comparisons of effects of PVFT programs are still being analyzed and reanalyzed, but one conclusion which usually appears in these reports (e.g. Stallings, 1973; Stallings and Kascovitz, 1974) is that the structured programs, such as the Pittsburgh and the Kansas model, yield greater pupil achievement in reading and math than do the more flexible programs. The likely cause for this superiority would appear to be the materials and the instructions for their use. It seems unlikely--at least at this time--that the different PVFT programs differ in other variables such as teacher expectations, teacher selection, teacher abilities, or teacher training, and the conclusion that I currently prefer is that the differences in program effects in reading and math are due to the types of materials used and the instructions for their use. Such a conclusion about the facilitative effect of <u>some</u> curriculum programs needs to be tested and modified, and one purpose of this paper is to encourage such testing and modification.

One implication of this idea is that more time might be devoted to evaluating and testing the instructional materials which are chosen and used in classrooms. Such materials include not only well organized curriculum packages, but textbooks as well.

Identifying Curriculum Elements

If some curriculums are more effective than others for similar ends, then learning why they are more effective might facilitate both curriculum development and curriculum modification. Presumably differences in content and emphasis will be part of the answer, but what these other elements are remains to be determined. Some possible variables are:

The organization of the materials for study

The questions which are provided for the teacher (e.g. provided in workbooks on study sheets for the teacher)

The feedback which materials provide for students.

Although apparently there has not been much work on the identification of the functions which materials play in facilitating instruction, the broad outlines of possible functional variables are present in any textbook on instruction. These functions include structuring learning, presenting material to be learned, soliciting responses, reinforcing and correcting, extending, promoting practice, maintaining assessing pupil learning, organizing and sequencing instruction, and providing for consolidation of learning.

If curriculum materials were analyzed in terms of the above list or a similar list, one might be able to suggest that some packages provide more tools for learning than do other packages. As an example, consider how materials can provide feedback to pupils. If pupils who are attempting to improve their reading comprehension use materials which are accompanied by multiple choice questions, they may obtain more feedback on their ability to understand a story than they would obtain through a class discussion. Furthermore, if the multiple choice questions were consistently organized by type so that a pupil learned whether he was having particular trouble figuring out words in context or getting the main idea of a paragraph, then such organization may provide more feedback than multiple choice questions alone.

In contrast to the usual emphasis upon coding interaction, the emphasis in the above examples is upon materials and determining whether

or how such materials help both the teacher and the pupil. Current observational procedures may obscure the usefulness of materials. For example, coding the types of questions which teachers ask may obscure the possibility that the teachers had to develop their own questions in one curriculum and were provided with both specific and illustrative questions in another curriculum. Similarly, the organization and maintenance of student practice may be well developed through structured materials in one curriculum and left up to the teacher's devising in another. Remediation exercises for pupils may be built into short, instructional packages in one curriculum, whereas in another, the teacher may be responsible for developing and organizing remediation materials.

The above examples are merely suggestive of research. Of course, one would not expect that all multiple choice questions in all curriculum packages are of equal merit, and the effects of different types of written questions has long been a topic of study. The point, however, is that the usual studies of classroom interaction have tended to focus only upon the questions which teachers or pupils asked and have not included analysis of the kind, sequence, and format of questions which the materials asked pupils. Whether such additional research into instructional elements present in the materials will be functional is unknown at present. My argument, however, is that the instructional variables inherent in materials should be included in the study of classroom learning.

Using Curriculum in Process-Outcome Research

This section contains some suggestions and issues in including knowledge of curriculum in process-outcome research. The ideas were developed from an inspection of current practices and are expanded and modified from an earlier piece on this topic (Rosenshine and Furst, 1973).

The first question might be the extent to which knowledge of curriculum predicted pupil achievement. That is, are there some packages or combinations of packages which are more effective in reaching the criteria than others. If the question is answered positively, how does

one then look at teacher behavior? Does one look at teacher behavior across curricula or does one only look within curricula?

There are at least three types of questions one can ask about different curriculum programs:

1. Are some curriculum packages (or combinations of packages) more effective than others on reaching the same goals?
2. If some programs are differentially effective, then are there general rules about effective elements <u>across</u> different packages? Are there some teacher behaviors which correlate with pupil outcomes across packages? Are there elements in the curriculum materials which are associated with greater pupil achievement? Some elements which might be considered are among those suggested in the previous section:

   The questions which are offered for the teacher to ask
   The organization of the materials for study
   The feedback which materials provide to students.
3. What are correlates and what are causes of outcomes within <u>individual</u> programs? What modifications in materials, organization and use of materials, and teacher behavior might facilitate the effectiveness of individual programs?

1. Comparison of outcomes

There are a number of possible steps in studying whether these different programs and combinations of programs are yielding different effects. One procedure--which was used by Soar (1971) to study processes but which seems applicable to outcomes also--would be to group the programs by type and use something like a multiple-range test to determine which groupings of programs are similar as to outcomes. When programs do not differ as to outcomes, then it would seem fruitful to use knowledge of curriculum in the subsequent analyses.

If programs were found to differ in effects, two additional analyses on effects seem relevant to aid in interpreting these results. The first analysis would look at effect by level of implementation.

Effects and level of implementation. A major question within the study of any curriculum package is the relationship between curriculum effects and level of implementation. Level of implementation refers to whether a program or a package is being implemented according to the specifications of the developers. The research question is whether levels of implementation (or levels of different implementation variables) affect the outcomes. That is, for some programs the desired effects may not occur unless the program is well-implemented, whereas other programs may be so robust that when poorly implemented they may be as effective as when well-implemented. On one hand, then, it can be argued that it doesn't make sense to compare the effects of well-implemented versions of program A against poorly implemented versions of program B. On the other hand, it is also possible that for program B the implementation variables are not so important as we thought and that a version labeled as poorly implemented is as effective as a version labeled as well-implemented.

In any discussion of the effects of curriculum packages, one has to note that in the First Year Reading Study (Bond and Dykstra, 1967) in which a variety of reading approaches and materials for teaching these approaches were compared, the general result was that one approach or set of materials was not consistently superior to another. A number of possible explanations for these results should be considered. One explanation could be that the curriculum programs did not differ between themselves in content covered or content emphasized, and this lack of difference was reflected in the results. Another possibility is that the results cannot be adequately interpreted without information on program implementation. Some programs may have been more easy to implement than others, and degree of implementation may have had different importance in different programs. It is thus suggested that without data on implementation (and on the importance of implementation) comparisons between programs may be premature.

Implementation may be a critical variable in some programs and not in others. In one report (Stallings and Kaskovitz, 1974), correlations were computed between the total implementation score in each of the classrooms within each of the Planned Variation Follow Through (PVFT) and the

adjusted measures of pupil gain in reading and mathematics. The results did not follow any pattern. In most cases the correlations were low (.0 to .2), and in those cases in which results were .3 or better, the overall pattern was not consistent across grades or subject area within the FT model being studied.

Information on level of implementation can be used a number of ways. One procedure would be to limit comparisons of outcomes to those classrooms which are labeled as high implementers. A subsequent analysis can determine whether level of implementation is an important variable in different programs. In cases in which level of implementation is not as important as was thought, then all classrooms using a program can be included in the analysis.

(Research using level of implementation variables will continue to be difficult because in addition to studying the effects of overall measures of implementation one would also need to study the effects of individual implementation variables.)

2. Determining correlations across programs.

In studies within Planned Variation Follow Through (Soar, 1971; Stallings and Kaskovitz, 1974) correlations between observed behaviors and pupil outcomes were computed across the seven or eight programs without regard for the curriculum in which the behaviors took place. Thus, variables such as the kind of questions teachers asked, the amount of time pupils spent on academic work, the amount of initiation pupils did, and the types of responses teachers make to pupils were coded and the frequencies of these variables were correlated with the adjusted outcome measures without any consideration of the program in which these took place.

The authors of these reports have interpreted such results as giving general information on the relationship of the process variables and student outcomes. Although I would like to accept their results as such, there are several questions and problems which are still unanswered. First, if the programs are being well implemented, then teacher behaviors

within programs should vary very little. That is, one would expect that the teacher behavior of elaborating on a pupil initiated idea would be infrequent in the structured programs and fairly frequent in the flexible programs with little variation within classrooms within a structured or a flexible program. If there is a strong correlation between teacher elaboration and pupil outcomes, such a correlation may represent differences in programs and not independent behavioral variables which are related to pupil outcomes. Thus, once again there is an implication of instances in which it is the curriculum and not the instruction per se that makes a difference.

3. Determining within program correlations.

Research within curriculum simply involves searching for correlates of pupil differential outcomes associated with a particular curriculum or instructional package. Such a search could take place by manipulating potentially important instructional variables and noting the effects of the manipulation, or it could take place by observing natural variation in instruction which occurs as a curriculum is being implemented and ascertaining which, if any, elements of this natural variation are related to the outcomes. Hopefully the two processes feed into each other, with the results of the experimental studies suggesting new correlational variables, and the results of the correlational studies suggesting new experimental variables.

In the unresolved area of how to analyze and interpret results across programs, it would seem appropriate to conduct more research on correlates and causes of effectiveness within programs. Suppose that some teachers were using McGraw Hill materials for teaching reading comprehension, while other teachers were using the Barnell Loft Specific Skill Series and still others were using the Science Research Associates (SRA) reading boxes. Even if one identified correlates of outcomes across programs, one would not know if such variables operated as correlates within these different programs. Because one wishes to apply results to specific programs, it would seem relevant to study how these variables operate within specific programs.

If the curricula yield different effects, then one can wonder whether it worth continuing to look at teacher behaviors across curriculums. That is, if classrooms using the Specific Skills Series attained higher reading achievement than was obtained in other packages, looking at differential teacher interactive behavior may be inappropriate because in this series most of the work is done individually and children are able to check their answers quickly and move to remedial work or higher level work. In such a structured situation, there is minimal need for teacher-pupil interactions, and therefore it is likely that there would be fewer displays of teacher presentations, teacher enthusiasm, or higher order questioning. The more appropriate teacher behaviors would appear to be more managerial than presentational.

At the same time, how does one look at a teacher's assignment of remedial work if we know that some classrooms are using the Specific Skill Series? It is possible that the behaviors of assigning remedial work will not differ greatly in frequency across curriculum programs, but such assignment may mean something different when it takes place within the context of the Specific Skill Series.

Thus, regardless of how one feels about the importance of accepting curriculum as a contextual variable in the general study of the effects of teaching, it would seem important to study teacher behaviors and their effects within curriculum packages. Although some research has been conducted within programs (e.g. Kochendorfer, 1967; Siegal and Rosenshine, 1972; Siegal, 1973) there has been relatively little research of this type, despite its apparent utility for improving specific programs.

The Relevance of Content Covered and Content Emphasized

The relevance of content covered in assessing curriculum effects is an area which needs much more research. On one side, it can be argued that what is taught is more important than how it is taught, and evidence on this point will be presented in this section. But if this argument is pushed, one could claim that the results of any curriculum package cannot be compared with another package because each package is emphasizing different criteria.

Program content and program emphasis. Based on their comparison of results of twenty-three studies which compared innovative and traditional curriculums, Walker and Schaffarick (1974) concluded that different curriculums produced different patterns of achievement, and suggested that these patterns are strongly influenced by the content that is included and emphasized within each. They assert that because different curriculums in the same subject area emphasize different content and skills to be learned, one does not gain much from comparing these curriculums. Furthermore, they argue that against such powerful variables as content and emphasis, other variations such as teaching procedures have relatively less influence on outcomes. They are, of course, aware of the limits of their conclusion. "We would be foolish to expect, for example, that simply including more content in texts and tests would ensure ever greater knowledge and skill." And, they are aware that their conclusion exists in the aggregate and is limited to patterns of achievement in large groups of students studying roughly similar kinds and amounts of content.

To be sure, they present their claim as an idea to be tested rather than as an accomplished fact. They suggest, furthermore, that different sorts of presentations of an identical item of content will no doubt produce differences in achievement and recommend further research to determine why certain ingredients contribute to a successful curriculum presentation.

One approach to the question of effective teaching might be to determine first the influence of three variables: content covered and

emphasized, time available for instruction, and student attention to task. The hypothesis would be that these variables are more powerful than any instructional variables. If the above hypothesis can be demonstrated, the next step would be what other variables, if any, have impact.

International variations in content covered. It is also possible to suggest the effects of variations in content in different countries. Perhaps the most dramatic study of the importance of content was conducted by Pidgeon (1970). In this study, the mathematics achievement of 3,000 eleven-year-old pupils in California was compared with that of a similarly stratified sample of 3,000 eleven-year-old pupils in England. Two thousand English pupils achieved a score above 35 on the 75 item test, whereas only 54 (sic) California pupils achieved a score above 35. Pidgeon (1970) inspected the mathematics textbooks used in the two countries and concluded that the material in the material in the English textbooks was one to two years in advance of the material in the California textbooks. He concluded that the English pupils were able to learn more mathematics because the teachers expected more learning from them, and these teacher expectations were reflected in the content that was covered.

Additional evidence on the importance of content comes from two studies conducted by the International Association for the Evaluation of Educational Achievement (IEA). In two cross-national studies, one in mathematics and one in science, the teachers were asked to look at the items on the test and indicate the percentage of students in the class who had the opportunity to learn the material covered by the item. When these scores were assembled by country there was a strong positive correlation between opportunity to learn material and actual achievement scores by country. (When correlations were computed between opportunity to learn in individual schools and student achievement within country the correlations were not significant, possibly because of limited variation within each country). One could argue

that the content which is covered indicates the national expectancy.

Within course research. It is also possible to note the importance of content and emphasis upon learning the content by looking at studies which took place within a curriculum or within a course of study. Rosenshine (1968); Shutes (1970); and Chang and Raths (1971) all found significant relationships between an assessment of the content or skills covered and pupil achievement. Stallings and Kaskovitz (1974) found that class activities involving numbers, number of children involved in mathematics, and total verbal interactions regarding math were all significantly and positively correlated with math gain in both the first and third grade; number of activities involving reading and number of children involved in reading was significantly and positively correlated with reading gain in the first grade but not in the third grade.

As part of their two-year study of process-outcome relationships in 31 first and 28 second grade classrooms, Brophy and Evertson (1974) asked observers to estimate the percentage of time that teachers devoted to various curriculum areas. Although some positive correlations were obtained, the correlations between time in math, or time in reading, and appropriate measures of pupil gain were not consistent or strong across each of the two years of their study. However, in their summary of all their results, Brophy and Evertson concluded that the teachers who had the most success were those who were most determined to teach the students, who took responsibility for making sure that pupils learn, and who were determined to overcome instructional obstacles. The teachers who got the greatest gains did so by "overteaching," by presenting material in small doses, with greater repetition and opportunity for practice than less successful teachers.

Another approach to the importance of content and teacher ability to cover appropriate content might be derived from the study by Pellegrini and Hicks (1972). In this study, tutors worked individually with elementary school children. Three groups provided comparisons to

assess the expectancy treatment: One group of tutors was told their tutees were of high ability and would make dramatic gains in academic areas in the next few months; a second group was told their students were of average ability; and a third group was told their children were below average in ability. After 17 weeks of tutoring, the children were retested on the PPVT and the WICS Similiarities Test. There were no significant differences between these three treatment groups. A fourth group was also told that their students were of high ability, and this set of tutors was shown sample items from PPVT and WICS Similarities Test. This group scored signfiicantly higher than the other three on the PPVT probably because the tutors' familiarity with the instrument resulted in instruction in the discrete skills necessary for successful test performance. While this group did not perform better than the others on the WISC, the tutors probably did not possess sufficient skills to teach the more complex reasoning required by items on this test.

A series of studies involving trainees tutoring pupils for periods of thirty minutes or less have also yielded mixed results. In these studies (Beez, 1968; Carter, 1969; Panda and Guskin, 1970; and Brown, 1970) the tutors were given fictitious information about their tutees which was designed to induce a set for a low achieving or a high achieving pupil. In three of the studies (Beez, 1968; Carter, 1969; and Brown, 1970) the teachers attempted to teach more words to the "high achieving pupils." In the fourth study (Panda and Guskin, 1970) there was no difference in the number of words taught. The results on the number of words learned were fairly consistent; students learned more words in two (Beez, 1968; Carter, 1969) of the three studies in which they were taught more words.

In summary, the results of some of the above studies suggest that content covered, opportunity to learn, and a teacher's emphasis upon student achievement are important variables for continued study. One wishes that there were more studies on this topic and that the results of the above studies were stronger.

What is taught and how it is taught. If content covered is as important as the results appear to indicate, it seems worthwhile to probe the limits of a proposition such as "what is taught is more important than how it is taught." Walker and Schafferick suggest that one cannot expect that simply including more content in texts and tests would ensure greater knowledge and skill. Furthermore, we have learned from hundreds of experimental studies that different forms of presentation of identical content have frequently led to differences in achievement.

Although the results on content covered are strong, these correlations are seldom above .5 and the results are not always significant or consistent. Overall content covered appears to be an extremely important variable which has not received sufficient attention in research on teaching effects. Furthermore, it remains to be determined which procedures for teaching identical content and skills are most effective. However, it is suggested that previous research has been overly concerned with how something is taught and has not spent enough time identifying what has been taught. Perhaps future work based on observation of different classrooms could first establish what has been taught, then look at how it has been taught, and then look at the interactions between them.

One place where one could probe whether knowledge of how something is taught adds to what is taught might be through inspection of the results of the Planned Variation Follow Through research. The programs in PVFT range from strongly structured programs such as DISTAR to different forms of flexible and open classrooms. The results of these programs are not final and one expects that conclusions will be revised as new data come in and old data are interpreted, but the results to date are that the more structured programs are consistently obtaining greater pupil achievement in reading and mathematics than the more flexible programs. Assuming for the moment that these results will hold up (and although there has been consistent evidence favoring structured, high content emphasis programs for low income children one

should want to inspect the characteristics of isolated flexible classrooms which are equally strong in producing reading and math achievement), there are at least two possible reasons for the greater effectiveness of the structured programs:

1. The structured programs are emphasizing different ends (or different patterns of ends) from the flexible programs, or
2. Given similar ends, the materials and instructions for using the materials in the structured programs are superior to the materials and instructions in the flexible programs.

Each of these two possibilities might be probed in future research.

Probing the importance of content covered. One conclusion from the research cited above is that when instructional programs differ in the content which is presented and emphasized, these differences are reflected in student achievement scores. Furthermore, within an instructional program, variations in content covered are frequently positively and significantly correlated with student achievement, although these correlations are neither as large or as consistent as comparisons made across different programs.

Walker and Schaffarick (1974) have noted that although the above conclusions appear obvious, such obviousness is not reflected in decisions about curriculum packages and textbooks. Furthermore, the obviousness of considering content has not been reflected in the analysis of process-outcome studies or in the items on teacher-evaluation checklists.

Other content variables. There are two other variables which would seem related to content covered, variables which are covered in Carroll's (1963) model of school learning but which receive insufficient attention in research programs. One such variable is the amount of attention a pupil gives to the task. It would seem worthwhile to conduct more research on the contributions of these variables to student learning.

Another approach to the importance of content covered is suggested by Dubin and Taveggia (1968) in their review of studies comparing the effects of different methods of college teaching. The teaching methods compared included lecturing, discussion, supervised independent study, lecture-discussion, and unsupervised independent study. Dubin and Taveggia concluded that these different instructional methods did not produce differences in student achievement.

In discussing the results of their review, Dubin and Taveggia suggest that research among comparative teaching methods should first focus on the question "What is there that is the same about two teaching methods?" They suggest that the textbook(s) utilized is an outstanding commonalty in classes taught by different methods and that various methods of instruction have only minor impact compared to the impact of instruction. They further recommend that future studies should examine differences among textbooks rather than among teaching methods.

I interpret their conclusion as saying that when students have access to the same content and know what they are expected to learn (and when they are capable of learning from the textbook) then differences in instruction will have relatively little impact.

Eliminating Alternative Hypotheses

A number of those who have been reviewing research on teaching have suggested that, although the results to date have not been as strong as we hoped, results will be better if future research incorporates their latest ideas (Brophy and Good, 1974; Good and Brophy, 1973; Duncan and Biddle, 1974; Rosenshine and Furst, 1971, 1973).

The number of suggestions is extremely large. One wishes that we would spend more time eliminating unproductive ideas. I would hope that future studies could begin by explicitly stating why the author hoped that the results of his/her study would be stronger than the results of previous studies, and that the conclusion of the study would evaluate the merit of the author's new ideas. Unfortunately,

if a new idea is not validated by the study, the tendency is to fault the circumstances of the study and retain the original idea without modification.

One of the areas for which there are a number of suggestions is in how to code classroom transactions. Comments have been made as to the coding unit (e.g. time unit, natural break in purpose, sign unit), the number of co-variables to be included with an interaction (e.g. content, speaker, level of conceptualization, number of students attending), whether previous events should be included in the coding unit, and the size of the variable in the unit (e.g. should instances of praise be subdivided as to type). One wishes that when someone suggests these new approaches (and, belatedly, I include myself) he would also discuss the types of results which would indicate that the new approach does <u>not</u> have merit.

LIST OF REFERENCES

Beez, W. Influence of biased psychological reports on teacher behavior and pupil performance. Proceedings of the 76th Annual Convention of the American Psychological Association, 1968, 3, 605-505.

Bond, G.L. and Dykstra, R. The cooperative research program in first grade reading instruction. Reading Research Quarterly, 1967, 2, 1-142.

Brophy, J. and Evertson, C. Process-product correlations in the Texas Teacher Effectiveness Study: Final Report. The University of Texas at Austin, 1974.

Brophy, J.E. and Good, T.L. Teacher-student relationships: Causes and Consequences. New York: Holt, Rinehart, and Winston, 1974.

Brown, W. The influence of student information on the formulation of teacher expectancy. Doctoral dissertation, Indiana University, 1970.

Carroll, J.B. A model of school learning. Teachers College Record, 1963, 64, 723-732.

Carter, R. Locus of control and teacher expectancy as related to achievement of young school children. Doctoral dissertation, Indiana University, 1969.

Chang, S.S. and Raths, J. The schools' contribution to the cumulating deficit. Journal of Educational Research, 1971, 64, 272-276.

Duncan, M. and Biddle, B. The study of teaching. Holt, Rinehart, and Winston, 1974.

Evans, E.D. Contemporary influences in early childhood education. New York: Holt, Rinehart, and Winston, 1971.

Good, T. and Brophy, J. Looking in Classrooms. New York: Harper and Row, 1973.

Panda, K. and Guskin, S. Effect of social reinforcement, locus of control, and cognitive style on concept learning among retarded children. Center for Educational Research and Development for Handicapped Children, Indiana University, 1970.

Pellegrini, R. and Hicks, R. Prophecy effects and tutorial instruction for the disadvantaged child. American Educational Research Journal, 1972, 9, 413-419.

Pidgeon, D. Expectation and pupil performance. Slough, Great Britain: National Federation for Educational Research, 1970.

Rosenshine, B. Objectively measured behavioral predictors of effectiveness in explaining. In I. Westbury and A. Bellack (Eds.) Research into classroom processes. New York: Teachers' College Press, 1971. Pp. 51-98.

Rosenshine, B. and Furst, N.F. Research on teacher performance criteria. In B.O. Smith (Ed.) Research in teacher education. Englewood Cliffs, N.J.: Prentice Hall, 1971.

Rosenshine, B. and Furst, N.F. The use of direct observation to study teaching. In R.M.W. Travers (Ed.) Second Handbook of Research on Teaching. Rand McNally, 1973.

Rosenshine, B. New directions for research on teaching. In How Teachers Make a Difference. U.S. Government Printing Office, 1971.

Shutes, R.E. Verbal behaviors and instructional effectiveness. Doctoral dissertation, Stanford University. University Microfilms 70-1063. 1970.

Siegel, M. An experimental investigation of teacher behavior and student achievement in the DISTAR program. Paper presented to the American Educational Research Association, 1974. (School of Education, University of Illinois at Urbana).

Siegel, M. and Rosenshine, B. Teacher behavior and student achievement in the DISTAR program. Paper presented to the American Educational Research Association, 1973. (School of Education, University of Illinois at Urbana).

Soar, R.S. An empirical analysis of selected Follow Through programs. In I.J. Gordon (Ed.) Early Childhood Education. Chicago: University of Chicago Press and National Society for the Study of Education, 1972.

Stallings, J. Follow-Through program classroom observation evaluation 1971-1972. Menlo Park, Ca.: Stanford Research Institute, 1973.

Stallings, J. and Kaskowitz, D. Follow-Through program classroom observation evaluation 1972-1973. Menlo Park, Ca.: Stanford Research Institute, 1974.

Walker, D. and Schaffarick, J. Comparing curricula. Review of Educational Research, 1974, 44, 83-111.