DOCUMENT RESUME

ED 231 794 SP 022 666

AUTHOR	Slavin, Robert E.
TITLE	Component Building: A Strategy for Research-Based Instructional Improvement. Report Number 337.
INSTITUTION	Johns Hopkins Univ., Baltimore, Md. Center for Social Organization of Schools.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	May 83
GRANT	NIE-G-83-0002
NOTE	36p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; *Change Strategies; Class Organization; *Classroom Research; Conceptual Tempo; Course Organization; Decision Making; Elementary Secondary Education; *Instructional Design; *Instructional Improvement; Programed Instruction; *Research Utilization; Student Motivation; Teaching Methods; Time Factors (Learning)

ABSTRACT

Component-building research on practical issues of instructional design can make a substantial contribution to research-based school improvement. Component-building research consists of rigorous within-school experiments that test various components of instruction for effects on student achievement. Research would address practical issues that teachers and principals face each year--for example, should homework be given every day, should films be used more or less often, should a school use a reward-for-effort or continuous progress grading system, should students be regrouped for mathematics instruction, and so on. A large number of such rigorous studies conducted by teachers, principals, and researchers could produce a science of instructional design. Complete programs of effective instruction could be developed by combining components that are proved to increase student achievement. (Author/JD)

Center for Social Organization of Schools

ED231794

REPORT NUMBER 337

MAY 1983

COMPONENT BUILDING: A STRATEGY FOR RESEARCH-BASED INSTRUCTIONAL IMPROVEMENT

Robert E. Slavin

STAFF

Edward L. McDill, Co-Director

James M. McPartland, Co-Director

Karl L. Alexander

Henry J. Becker

Jomills H. Braddock, II

Shirley Brown

Ruth H. Carter

Michael Cook

Robert L. Crain

Doris R. Entwisle

Joyce L. Epstein

James Fennessey

Denise C. Gottfredson

Gary D. Gottfredson

Linda S. Gottfredson

Edward J. Harsch

John H. Hollifield

Barbara J. Hucksoll

Ramona M. Humphrey

Lois G. Hybl

Helene M. Kapinos

Nancy L. Karweit

Hazel G. Kennedy

Marshall B. Leavey

Gretchen M. Luebbe

Nancy A. Madden

Kirk Nabors

Alejandro Portes

Donald C. Rickert, Jr.

Laura Hersh Salganik

Robert E. Slavin

Jane St. John

Valarie Sunderland

Gail E. Thomas

William T. Trent

COMPONENT BUILDING: A STRATEGY FOR RESEARCH-BASED INSTRUCTIONAL IMPROVEMENT

Grant No. NIE-G-83-0002

Robert E. Slavin

Report No. 337

May 1983

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States National Institute of Education, Department of Education. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the Institute should be inferred.

Center for Social Organization of Schools
The Johns Hopkins University
3505 North Charles Street
Baltimore, MD 21218

The Center

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through three research programs to achieve its objectives. The School Organization Program investigates how school and classroom organization affects student learning and other outcomes. Current studies focus on parental involvement, microcomputers, use of time in schools, cooperative learning, and other organizational factors. The Education and Work Program examines the relationship between schooling and students' later-life occupational and educational success. Current projects include studies of the competencies required in the workplace, the sources of training and experience that lead to employment, college students' major field choices, and employment of urban minority youth. The Schools and Delinquency Program researches the problem of crime, violence, vandalism, and disorder in schools and the role that schools play in delinquency. Ongoing studies address the need to develop a strong theory of delinquent behavior while examining school effects on delinquency and evaluating delinquency prevention programs in and outside of schools.

The Center also supports a Fellowships in Education Research program that provides opportunities for talented young researchers to conduct and publish significant research and encourages the participation of women and minorities in research on education.

This report, prepared by the School Organization Program, examines how educational research contributes to instructional improvement. The report advocates a "component-building" program of research consisting of field experiments on classroom practices that are separate in themselves, but are or could become components of complete programs.

Abstract

This paper proposes that component-building research on practical issues of instructional design can make a substantial contribution to research-based school improvement.

Component-building research consists of rigorous within-school experiments that test various components of instruction for effects on student achievement. The experiments would address the practical issues that teachers and principals face each year -- for example, should homework be given every day, should films be used more or less often, should a school use a reward-for-effort or continuous progress grading system, should students be regrouped for math instruction, and so on.

A large number of such rigorous studies conducted by teachers, principals, and researchers could produce a science of instructional design. Complete programs of effective instruction could be developed by combining components that are proved to increase student achievement.

Acknowledgment

The author would like to thank Nancy Madden, Gary Gottfredson, Nancy Karweit, William Trent, Thomas Good, Barak Rosenshine, and Mary Rohrkemper for their comments on this paper.

Component Building:
A Strategy for Research-Based Instructional Improvement

Because education is an applied field, one important task of educational researchers is to conduct research directed at identifying specific instructional practices that teachers may use to increase their students' achievement. Researchers are often asked questions about many of the basic instructional decisions teachers must make in preparing a lesson or course. For example, researchers might be asked whether it is better (in a particular subject) to teach a single lesson to the entire class, to teach different lessons to homogeneous subgroups, or to give students individualized materials (such as programmed instructional materials) to complete at their own rates. Is it more conducive to achievement to allow students to work with each other or to forbid them to do so? Should students be graded on a curve, against preset standards, based on improvement or effort, or should they not be graded at all? How frequently should quizzes and tests be given? What sort of questionning strategies and corrective feedback should teachers use? What is the best pace of instruction, and how should pace be set? How important is regularly assigned homework? How much choice should students have about what they study and how they do so? Is it better to challenge students with difficult problems or to let them experience frequent success by giving them easy problems?

Researchers and others interested in improvement of instructional methods must consider these and other questions when they set out to create new instructional programs, to look for effective practices in current use, or to make recommendations to teachers about how they should teach to maximize student achievement. These questions relate to what can be called the <u>alterable</u>

elements of classroom instruction; they are all practices or policies that are (at least in principle) within the span of control of individual teachers. There are also many school-level or district-level policies or practices that may impact on student achievement, and school- and district-level decisions (such as allocation of resources and staff, grading or promotion standards, bussing or mainstreaming policies, and so on) certainly may expand or contract the range of practicable alternatives for the teacher. However, it is useful to restrict attention for the moment to practices available to the classroom teacher, and to ask whether we can build a science of classroom instruction based on a few broad assumptions about the characteristics and resources of the vast majority of elementary and secondary classrooms: twenty to forty students with moderate to extreme heterogeneity in preparedness, ability, and experience in the subject at hand; one teacher, perhaps (rarely) with an aide or regular volunteers; reasonably adequate books and other curriculum materials and supplies; and minimal to moderate discretionary funds for special materials or activities.

Assuming even the most minimal level of resources, the range of alternative means of organizing the classroom for instruction is quite wide. However, the consequences of choosing each option are poorly understood and probably quite complex, as they may interact with other features of classroom organization and student, teacher, and school characteristics. For example, an optimum instructional pace might be different in a heterogeneous class than in a homogeneous one, might be different in a class using some form of individualization than in one using whole class instruction, and might be different in a secondary mathematics class than in an elementary social studies class. Also, learning is not the only important outcome of schooling. If it turned out that caning or public ridicule of low achievers were an effective means of motivating students to learn, we would (we should) still reject these stragegies because they would not create the kind of setting in which we would like to have children socialized. An instructional method that improved students' self-esteem, behavior, or race relations might be justifiably used if it had no negative effects on student achievement as compared to traditional or state-of-the-art instructional methods.

Alterable Elements of Classroom Organization

However, restricting our attention to learning outcomes as the measure of instructional effectiveness, it is useful to have an overall framework of alterable elements of classroom organization to conceptualize the many alternatives (discussed above) for presenting instruction. Carroll (1963), in his classic "model of school learning," proposed that five elements were involved in predicting students' rates of learning. These were as follows (Carroll, 1963, p. 729):

1. Aptitude. The amount of time needed to learn the task under optimal instructional conditions;

2. Ability to understand instruction;

3. Perseverance. The amount of time the learner is willing to engage actively in learning;

4. Opportunity. Time allowed for learning; and

5. Quality of Instruction. The degree to which instruction is presented so that it will not require additional time for mastery beyond that required in view of aptitude.

Carroll's model, in particular its emphasis on the relationship between time spent learning to time needed to learn, has guided much of the study of instructional dimensions in the past decade. However, its usefulness for understanding the alterable elements of classroom organization is limited by its mix of elements internal to students (e.g., aptitude) which are difficult to affect in the short run, and elements external to students (e.g., quality of instruction) that are under the control of the teacher or school and are therefore alterable elements of instruction.

I would propose four critical dimensions of classroom organization as the <u>alterable</u> elements of classroom organization implied in Carroll's model.

1. <u>Quality of Instruction</u>. The degree to which the proper information or skills are presented to students in an appropriate form and sequence. This is directly from Carroll's model.

2. <u>Appropriate Level of Instruction</u>. The degree to which students have the prerequisite skills to understand the material presented to them (from Carroll's "ability to understand instruction") but have not already learned it. Obviously, if students have already learned what the teacher is teaching, or do not have the prerequisite skills to learn it, the instruction is ineffective in producing learning gains.

3. <u>Incentive</u>. The degree to which students are motivated to work on instructional tasks (Carroll's "persistance") and to retain what is taught.

4. <u>Time</u>. The degree to which students are given adequate time to learn what is taught (Carroll's "opportunity").

What is important about these four elements is that <u>all four</u> must be adequate if learning is to take place. For example, no amount of incentive to learn will matter if the material being presented is wrong or useless, if students do not have the prerequisite skills or have already learned the material, or if students do not have enough time to learn the material. If students are unmotivated to learn what is taught, the other elements will make no difference, and so on.

It is hardly revolutionary to suggest that students will learn appropriately presented material most efficiently if they are motivated, have the prerequisite skills, have not already learned what is being taught, and have adequate time in which to learn it. However, I would argue that some of the great disappointments in research relating to instructional improvement stem from a failure to simultaneously consider all four of these criteria.

For example, consider the case of programmed instruction. Programmed instruction was primarily developed to solve the problem of differences in student learning rates and levels of prior knowledge by allowing students to work on materials at their own levels and rates. Programmed instruction is thus probably very effective in assuring that students have the prerequisite skills for whatever they are learning and that they do not waste their time on material they already know. Students have as much time as they need to learn the material, so adequate time is not a problem. The programmed materials themselves are carefully constructed and piloted, so content should not be a problem. Despite these strengths, reviews of the effects of programmed instruction on student achievement (eg., Ebeling, 1981; Jamison, Suppes, & Wells, 1974; Miller, 1976; Thompson, 1975; Zoll, 1969) uniformly conclude that programmed instruction is rarely superior to traditional methods in promoting student achievement.

There are three probable flaws in programmed instruction that could account for these disappointing findings. One is in the area of incentive. Teachers report that programmed instruction can be boring for students (see, for example, Kepler and Randall, 1977). There is often little reason for students to want to make rapid progress in their programmed workbooks or to want to maintain a high degree of accuracy. Programmed instruction isolates students from one another and from the teacher; if a teacher used his or her time very efficiently during a fifty-minute period, he or she could visit with each student in a class of twenty-five an average of only two minutes per day, most of which would be required for checking materials rather than for teaching or encouragement. Yet record-keeping and other management activities require much teacher time, reducing time for teacher-student interaction still further. Teacher-student interaction is the source of some of the incentive to learn in traditionally structured classrooms that may be lacking in programmed instruction.

While programmed instruction may provide plentiful time for students to learn from their instructional materials, rates of learning may be reduced by the need for students to spend much time involved in management-related activities. If students are poorly motivated to work on their programmed materials, this may result in losses of time-on-task. Thus, programmed instruction may not create conditions for optimal use of time.

Another possible flaw of programmed instruction is in the area of quality of instruction. It may be that many students learn poorly from print materials alone, but need verbal instruction in addition to or instead of print materials. This is particularly likely to be true of poor readers or other low achievers. Programmed instruction may not provide adequate opportunities for corrective feedback, as when students need re-explanation of concepts they did not grasp from the programmed materials. The time needed for management-related activities may reduce or eliminate time for direct instruction, corrective feedback, or other concept-related (as opposed to management-related activities) interaction with a teacher.

Thus, even though programmed instruction seems to solve some of the most critical problems of group instruction, it apparently creates enough new problems to cancel out its benefits. What it gains in providing appropriate levels of instruction and plentiful, flexible time, it might lose in incentive, time lost to management-related activities, and quality of instruction.

There is no inherent reason that programmed instruction _must_ have the problems listed above. For example, there are at least two new programs that use programmed instructional materials but also use classroom organizational methods that are specifically designed to remedy the limitations of this approach. One is Wang's (1982) Adaptive Instruction program, which attempts to solve the (presumed) incentive problem by allowing students to choose their own schedule of activities and by interspersing high-interest activities with more routine work. It attempts to increase direct teaching time (and improve quality of instruction) by the use of aides and/or team teaching, and by having some students do work that requires little teacher direction while others work in small groups with the teacher. Another new programmed instruction method is the Team Assisted Individualization (TAI) program (Slavin, Leavey, & Madden, in press), in which students work in heterogenous teams on individualized materials. TAI confronts the incentive problems of programmed instruction by having students work toward team scores based on the number of units completed by all team members each week and the accuracy of their final tests. The teams who meet preset criteria receive certificates and special privileges. This team score system is hypothesized to create within-team pressures toward high achievement. TAI is designed to allow for more direct teaching time without requiring an aide or team teacher by having students check all of each others' work. The teacher meets with homogeneous math groups to provide direct instruction while the other students continue with their programmed instructional materials. The teacher is involved in the cycle of activities relating to the individualized materials only when students fail to pass any of several mastery checks, which is rare.

Probably because of their attention to motivation and direct instruction, both the Adaptive Instruction model (Wang, 1982) and the TAI model (Slavin, Leavey, & Madden, 1983, in press) have had positive effects on student achievement as compared to traditionally taught control groups. This research gives some indication that programmed instruction is not inherently unworkable or ineffective, but can be made workable and effective by the use of classroom organizational elements designed to enhance motivation and increase direct instructional time.

A more recent disappointment in educational research is research on time-on-task. It seems obvious that the more time students have to learn, and the more time they actually spend learning, the more they will achieve, and in fact studies of the relationship between time and learning do tend to show a positive correlation between these measures, controlling for prior achievement or ability. However, the positive relationship tends to be so small that it often fails to achieve statistical significance, much less substantive importance. Researchers have rarely found a significant effect on achievement on such gross measures of available time as the number of hours in the school day or days in the school year (see Frederick & Walberg, 1980; Karweit, 1976). Even using the more proximate measure of the number of minutes students actually spend on-task, the Beginning Teacher Evaluation Study (Marliave, Fisher, & Dishaw, 1978) found significant effects of time-on-task on achievement at the fifth grade level, but not at the second grade level. None of their analyses of the effect of engaged minutes on task accounted uniquely for more than one percent of the variance in total reading or math performance, controlling for pretests.<*> In a similar study, Karweit and Slavin (1981) found a significant effect of engaged minutes on mathematics achievement in grades 2-3, but not in grades 4-5, the reverse of the BTES findings; but again, no analyses uniquely accounted for more than two percent of the variance in posttest scores, controlling for pretests. The frequently cited Lahaderne (1968) study found statistically significant effects of attentiveness on achievement, partialling out IQ, in only three of eight analyses, and these effect sizes were small.

These weak and inconsistent effects were found even though the classes in these three studies (and others) varied markedly in time allocated to instruction, and even though student ability differences within classes, which are strongly related to engaged time, could not be completely controlled out, probably inflating the size of the findings. Between-class analyses, which are less contaminated by student ability differences and more focused on classroom

---

<*> There is a widespread belief that the BTES study did show strong effects of time-on-task on achievement (see, for example, Denham & Lieberman, 1980). However, the somewhat stronger effects usually reported for the BTES are for a composite measure called "Academic Learning Time," or ALT, which includes "percent of correct responses" to questions posed in class, largely a measure of student ability rather than a time-on-task measure. See Burstein (1980), Karweit (1982), and Karweit and Slavin (1981) for critiques of the BTES findings.

practices, were conducted in the BTES study (Marliave, Fisher, & Dishaw, 1978); no statistically significant effects of allocated time or engagement rate on achievement were found at either grade level.

How could engaged time have such a small impact on learning? The answer probably lies in the fact that studies of time-on-task rarely consider the other elements that I am arguing must be adequately implemented if learning is to take place; quality of instruction, appropriate level of instruction, and incentive. In particular, I would expect that the failure to find large effects of time-on-task on learning is due to a failure of the model linking these variables to adequately consider the role of appropriate level of instruction. When a teacher teaches some unit of information, there is an excellent chance that some proportion of the students in the class already know the information, and that others do not have the prerequisite skills to learn it in the time allowed. This is particularly likely to be problematic in such subjects as reading and math, where learning of most skills depends on mastery of prerequisite skills. If a teacher spends three periods teaching students how to divide two-digit numbers with remainders in the quotient, some students are likely to understand how to do the division in the first ten minutes (if they did not understand it before class). The rest of the instructional time is largely wasted for them. Other students, who perhaps never understood how to divide two digits <u>without</u> remainders in the quotient, will not be able to divide with remainders at the end of the three-period lesson. Their instructional time is thus also wasted. In fact, the only students for whom time would be strongly associated with learning are those who require 2-3 days (but no more) to learn the skill. However, averaging the effect of instructional time for such students with the (minimal) effect for rapid-mastery students and students who never master the material apparently leaves a very small positive effect of time on learning (see Karweit, 1978, 1982; Karweit & Slavin, 1981, for more on this topic).

It is not my purpose in this discussion to go into a detailed analysis of research on programmed instruction, time-on-task, or other topics. I am using these examples to illustrate the importance in research on classroom instruction of simultaneously considering all fou of the alterable elements of classroom organization outlined earlier in this paper: quality of instruction, a propriate levels of instruction, incentive, and time.

Traditional Approaches to Instructional Improvement: Master Teacher and Master Developer Research

Rosenshine (1982a) has made an interesting distinction between two current approaches to instructional improvement: the "master teacher" approach and the "master developer" approach. The "master teacher" approach, best represented by process-product research (see Brophy, 1979), is essentially a search for the methods used by teachers whose classes show the greatest gains in achievement over some period of time, often several successive years. These teachers' practices are observed, and compared to those of teachers whose classes make only average gains. The "master developer," on the other hand, creates a new instructional method out of whole cloth, usually based on well-established psychological or pedogogical principles. Examples of the products of such "master developers" are Mastery Learning (Block and Anderson, 1975), DISTAR (Becker & Carnine, 1980), the Missouri Mathematics Effectiveness Project (Good & Grouws, 1979), the First Grade Reading Program (Anderson, Evertson, & Brophy, 1979), Programmed Instruction (Glaser, 1965), the Exemplary Center for Reading Instruction (Rosenshine, 1982b), Adaptive Instruction (Wang, 1982), Student Team Learning (Slavin, 1980), and Team Assisted Individualization (Slavin, Leavey, & Madden, 1982). Each of these is a well thought-out, complete instructional program that is used in schools as such, not as a set of principles of instructional organization.

Where educational research has a direct impact on instructional practice, it is usually a result of either "master teacher" or "master developer" research. The impact of "master teacher" research usually shows up as a set of principles teachers try to follow, which they often encounter in teaching methods texts or other publications or from local in-service workshops, while the "master developer" programs are more typically implemented as complete programs that a district purchases or adopts. Sometimes "master developer" programs, such as the Good and Grouws (1979) and Anderson et al. (1979) programs, are based directly on the findings of "master teacher" (process-product) research, and it is often the case that general principles of effective instruction are drawn from such specific "master developer" programs as DISTAR (see, for example, Stallings and Kaskowitz, 1974). However, while "master teacher" and "master developer" strategies draw upon and enrich one another, their modes of inquiry and applications to instructional improvement are quite distinct.

Each of these approaches to research-based school improvement has its limitations. The "master teacher" research is inherently correlational, which means that we can never be sure that the "master teachers" are in fact good models. It is always possible that they have good students, who in turn make them look good. No amount of statistical control can, for example, completely isolate teacher practices associated with high achievement gains when some teachers have disruptive or unmotivated students while others have well-behaved, motivated ones. Some supposedly ineffective practices, such as spending large amounts of time "maintaining order," could be required in disruptive classes, while supposedly effective practices might be impossible in such classes. "Maintaining a low error rate" is another principle derived from process-product research that may reflect the fact that high achieving students tend to answer questions correctly rather than that asking easy questions increases student achievement.* Also, when we locate effective teachers, it is likely that we have located individuals who would do an excellent job implementing <u>any</u> instructional method. Since the "master teacher" research tends to produce evidence supporting traditional instructional practices, it could be argued that effective teachers are just the best implementers of a particular method (traditional instruction). Good implementers of any reasonable instructional method will almost certainly get better results than poor implementers of the same method. If all teachers used programmed instruction, process-product research would surely find that teachers who did a good job implementing these programs would get better results than teachers who did a less adequate job of implementing the programs (see Wang, 1982). However, this does not imply that programmed instruction per se is an optimal strategy. Similarly, process-product research should not be read to imply that traditional methods are optimal just because teachers who implement them best get the best results.

Perhaps most importantly, prescriptions for practice derived from "master teacher" research are inherently limited to the range of current widespread practice. It may well be that the most effective instructional strategies are

---

*In fact, Burstein (1980) reanalyzed the Beginning Teacher Evaluation Study data concerning this variable and found that when analyzed at the <u>class</u> level rather than the individual student level, low error rate tended to be <u>negatively</u> related to achievement gain.

not used at all or are not used enough to show up in the particular classrooms studied in this research. Few of the "master developer" programs would have ever been found already in existence in a teacher's repertoire (although most of the features of Good and Grouws' (1979) Missouri Mathematics Effectiveness program and Anderson, Evertson, and Brophy's (1979) First Grade Reading program were based on process-product research that identified practices used by effective teachers).

In contrast, research on the "master developer" programs is usually experimental, making direction of causality less of a concern, and these programs introduce methods that are not necessarily in current use, broadening the range of classroom instructional alternatives. However, the "master developer" approach to instructional improvement also has its limitations. One limitation is that as soon as a method is developed, usually at great expense and difficulty, the developer has a stake in proving its effectiveness. No matter how rigorous the developer is in evaluating his or her own program, many will be suspicious about such evaluations. However, this limitation is often only temporary, as when a program becomes widely known and used, independent evaluations are usually done by others.

Perhaps the most important limitation of the "master developer" approaches is that these programs are often prematurely solidified, before the various components of the programs have been independently studied. A new instructional method is typically developed, piloted, revised, evaluated, and (let us assume) found to be effective. It is then disseminated as a complete program. There is nothing wrong with this; if a program is effective as a complete package, it is appropriate to disseminate it as such. However, this progression may add little to a science of classroom organization, as all "master developer" programs contain many components, any one or combination of which could explain the effectiveness of the entire program. Some components of any of these programs are probably useless. Some may even be counter-productive. Since we do not know the effects of each component as such, we are often unable to adapt the programs to new uses, to improve them, or to learn from them in a way that would spark other development or research.

Another Approach to Instructional Improvement: Component Building

I would propose a third approach to research-based instructional improvement, which I will call "component building." What I mean by this is a long program of field experimental research on classroom practices that are or could become components of complete programs, but are separable elements in themselves. Some such components worth testing in randomized field experiments would be the principles derived from the "master teacher," or process-product research. Others might be components of the "master developer" programs, taken out and evaluated separately.

"Component-building" research might focus on the four alterable elements of classroom organization discussed earlier in this paper, or others proposed by researchers wiser than I. However, this need not mean that all of these elements must be included as factors in component-building research. Instead, programs designed to address these elements should be implemented, and then a single component could be evaluated by withholding it in some classes while using it in others. For example, let's take the Missouri Mathematics Effectiveness Program, or MMEP (Good and Grouws, 1979) as a basic model. The advantage of this particular model is that it is derived from the best of whole-class (i.e., not sub-grouped or individualized) traditional instruction, as most of its features are taken from process-product research comparing the most effective teachers to less effective ones (Good and Grouws, 1977). The Good and Grouws model emphasizes a high ratio of direct teaching to seatwork, an emphasis on meaning in mathematics instruction, controlled practice before seatwork, frequent assessment of student learning, checking of seatwork and homework, daily but brief homework assignments, weekly and monthly reviews, a rapid pace of instruction, and several class management strategies directed at increasing time-on-task. Each of these is an alterable component. Some are so sensible that they may not be worth investigation. For example, it is hard to argue against effective management strategies. However, many others would be interesting to study. For example, we might wish to train several teachers to use the MMEP, and then randomly assign some to use a rapid pace of instruction while others make sure that all students have mastered each skill before moving on. We might compare the MMEP with brief homework assignments to variations with no homework or long homework assignments, and so on. Alternatively, we might compare the MMEP to the same program with added components. For example, we might wish to know whether allowing students to work together during seatwork increases achievement more than having them work alone, or whether the use of advance organizers at the beginning of the lesson increases retention of the material taught, or whether a version of the MMEP that assigns students to ability-homogeneous math groups is more effective than the original program.

Of course, I would not propose the MMEP as the only "base" program for component-building research. Individualized programs, such as those of Wang (1982) or Slavin, Leavey, & Madden (in press) could be used as base programs for investigations directed at improving this type of instruction, and such subgrouped programs as ECRI (Rosenshine, 1982b) could be used as base programs for improving instruction involving homogeneous subgroups.

There are three primary advantages of this type of component-building research. First, by beginning with a well thought out, comprehensive program that has been shown to be effective, we can be relatively sure that many of the basic requirements of effective instruction (such as the four proposed above) are satisfied, at least to some degree. This may not be the case in traditional, untreated classrooms, where some components may appear to make no difference because something else is missing (for example, varying pace might appear to make no difference if students are unmotivated to learn or are receiving poor instruction). Second, evaluating components against a base of a comprehensive program solves an important problem of field experimental research: knowing what the untreated "control" teachers were really doing. Finally, the problem of Hawthrone effects are diminished by the fact that all teachers are using a new program.

Of course, I am not proposing that compoent-building become the only form of research directed at improvement of instructional methods. "Master developer" research must continue, both to provide new models to use as "bases" for component-building research and to crystallize new information about instructional effectiveness into comprehensive programs that schools can use now. "Master-teacher," or process-product research, must continue to suggest components worthy of study in field experiments. Laboratory research, including brief studies in the field, must also continue to investigate fine-grained issues difficult to evaluate in unavoidably less well-controlled, longer term field experiments, and also to propose additional components for further study. However, I would maintain that if an important goal of educational research is to provide teachers with more effective, practical instructional methods, component-building would be the primary route to this goal.

Examples of Component-Building Research To Be Done

Component-building research might best be organized around the alterable elements described above, or others proposed in the future. This section briefly describes a few investigations in each of these areas that might advance our understanding of the components of effective classroom instruction, as illustrations of the kind of of research I am proposing.

Research on quality of instruction might include studies of objectives-based teaching. For example, we might contrast programs in which teachers are given broad objectives to be achieved over the course of the year, special monthly objectives or unit objectives with pre- and posttests, or special weekly objectives with pre- and posttests and worksheets provided. Delivery modes, such as audio-visual, peer tutoring, and academic games and projects are components that have been extensively studied, as have such cognitive structuring strategies as use of advance organizers. One critically important issue of quality of instruction is departmentalization, which is being increasingly used at the elementary level. A presumed effect of departmentalization is to have teachers teach the subjects they are most competent to teach, but departmentalization may entail costs in terms of incentive (as teachers may have too many students to attend to the motivational needs of each one) and time (as students spend time changing classes). All of curriculum research, including the timing and sequencing of introducing various skills, falls under the area of quality of instruction.

Providing appropriate levels of instruction is a particularly important criterion of effective instruction in class groups that have a range of abilities or skills. Even if teachers had instantaneous information on how well students were learning from a whole-class lesson, they would still have great difficulties meeting divergent needs. Yet it is hard for teachers to accurately gauge the effects of their instruction on different students, making responding to divergent needs even more problematic. However, the alternatives to whole-class instruction also have serious costs in terms of instructional efficiency. If subgrouping is used (e.g., homogeneous reading or math groups), students may not effectively use their seatwork time, which is difficult for the teacher to monitor effectively. Since students in a class using three groups must spend at least two-thirds of their time working on their own, this is a serious problem. Tracking helps reduce student heterogeniety, but may create low-track classes that are difficult to teach. Pull-out programs for low achievers or for gifted students may help meet the needs of these students, but may also cause frequent disruptions, and coordinating instruction in the regular class with instruction delivered in resource rooms and gifted classes is very difficult. Corrective instruction given in class to students who do not initially achieve a mastery criterion may help low-achieving students, but may also waste the time of students who did achieve mastery. The problems of programmed instruction were discussed earlier. Thus, the issue of providing appropriate levels of instruction is complex, but each of the approaches discussed above must be studied as potential solutions to this problem. For example, if the problem with subgrouping is identified as the fact that students use their time not working with the teacher inefficiently, then special methods such as peer monitoring during that time might be evaluated. If lack of coordination between regular classes and resource rooms is a problem, then programs mandating such coordination must be evaluated, and so on.

Many problems of incentive stem from the fact that students enter class with different levels of skills, interest, and ability. Motivation to expend maximum effort is low when an individual has either a small chance of being rewarded regardless of effort or a good chance of being rewarded even if the individual expends relatively little effort (see Atkinson, 1958; Slavin, 1978). Yet traditional grading systems set up just such a situation. Low ability or low skilled students have very little chance to achieve high grades regardless of their effort, while high ability or high skilled students can often obtain high grades with minimal effort. Such incentive/evaluation methods as rewards-for-improvement, rewards-for-progress, and performance contracts are directed toward confronting this problem. Another potential problem of grading is that grades (and other feedback) are administered so infrequently that it is difficult for students to see the results of a change in effort or performance level; studies of frequency of performance feedback are needed. Also, grades may not be valued by many students. It may be important to reward school performance by tying it to more potent sources of rewards, such as parents (by having parents reward students based on regular school reports), peers (by rewarding students based on the performance of student learning teams), or non-classroom sources within the school (for example, making sports participation contingent on class reports).

It is possible that the grading system is not as important as day-to-day incentives or rewards. It may be very important how teachers call on students. Students might pay more attention in class or put more effort into classwork if they felt they had a good chance of being called on whether or not they raised their hand. Systems that insure that students will be recognized for knowing answers or doing good work but cannot hide if they do not do their best may be critical in motivating maximum effort. Formal point systems for in-class behavior and academic performance may be more effective than informal praise, but may have offsetting costs in terms of time to administer them.

Finally, classroom authority structures may relate to student motivation. If students have a choice in what they are to do or in what the class does, they may have more investment in the activity or outcome, but this benefit could be offset by the time it takes to administer a program with many choices and by the effect on quality of instruction of letting students have a role in deciding what they will study. For example, it may be harder to be sure that students learn math in a meaningful sequence if they can choose their math activities.

The total amount of time available for instruction is usually difficult to alter, but the way available time is organized can be varied. For example, instructional pace is an issue related to time. Is it more effective to teach many objectives per unit time or to teach more slowly to make sure that all students are mastering the lessons? Radical reorganizations of time use, such as intensive education (where students study one or two subjects intensively for several weeks instead of studying many subjects at the same time), should be studied, as should such simple interventions as eliminating all interruptions by banning use of the loudspeaker and forbidding other interruptions during instructional periods. Homework is a time-related issue. Assigning homework increases available time, but checking homework in class may (or may not) take too much valuable time from instruction to make it worthwhile.

Of course, many of these issues of quality of instruction, appropriate levels of instruction, incentive, and time have been extensively rsearched. However, experimental research with random assignment is still rarely seen in evaluations of most of these components, and until a component has been studied in well-controlled field experiments, we cannot be certain that we understand its effects on student achievement. Student ability is such a powerful variable that in correlational research, there is almost always a reasonable chance that practices presumed to increase achievement are instead results of a certain level of student ability. Differences between teachers also relate strongly to student achievement. No one would take seriously an educational experiment in which the experimenter allowed teachers to choose which experimental group to be in, or an experiment in which students in one experimental group started out much higher in achievement than another. No amount of statistical control can remove the effects of large differences between teachers in motivation or ability to use a particular program or large differences between students in prior achievement. Yet when we do correlational research on instructional programs or practices, we are usually comparing teachers or schools who chose one program or practice to those who chose another, and there are frequently important pre-existing achievement differences between students in schools that would choose one or another program or practice. For example, a correlational study of computer assisted instruction would be inherently flawed by the possibility that only schools that could afford computers and had a commitment to quality instruction would appear in the "experimental" group, and students in these schools might have done better than other students (even controlling for prior achievement) whether or not they had computers.

Even when problems of self-selection and pre-existing achievement differences are inconsequential, correlational research can point to teaching practices that may not be easily transferable to teachers not using them. For example, Kounin (1970) emphasizes "withitness" as a correlate of effective instruction. Can "withitness" be taught to teachers who are not "with it," and if so, will the newly "with it" teachers' students gain in achievement? Perhaps, but this is a question that can only be resolved in an experimental study.

Experimental Designs for Component-Building Research

To build a science of instructional design, we will need dozens, perhaps hundreds of field experiments in school settings examining the alterable elements of classroom organization. A glance through the major educational research journals would make one pessimistic about the chances of this occurring. Field experiments in schools are rare. Many researchers believe that rigorous field experiments simply cannot be done in schools, or at least cannot be done without massive resources. There are indeed formidable practical as well as experimental design related obstacles to overcome. However, high-quality, unbiased research can be done in schools without major expenditures. Some research designs suited to this purpose are described below (see also Campbell & Stanley, 1963; Kerlinger, 1979; Slavin, in press).

1. Random Selection of Teachers from among Volunteers. It is often possible to find a group of teachers to volunteer to participate in a research project, knowing that they have a 50-50 chance of being in the experimental (or control) group. Once a group of volunteers is identified, they can be randomly assigned (e.g., by coin flip) to experimental or control conditions. If at least four teachers are assigned to each treatment group, the chances that differences between teachers will obscure differences between treatments are slight. Pre-existing differences between classes can be controlled for and the power of the analysis can be increased by use of analysis of covariance or equivalent statistical procedures, with pretests, standardized test scores,

past grades, or other measures as covariates.

2. Random Selection of Classes Taught by the Same Teacher. In schools that use some degree of departmentalization, the same teachers might teach at least one experimental class and at least one control class. For example, if Ms. Calculation teaches two math classes, the researcher could flip a coin to determine which will be in the experimental group and which in the control group, repeating the procedure with several teachers. Analysis of covariance would again be used to control for minor class differences and increase statistical power.

3. Experimental Reversal Designs. Reasonably valid experiments can be done without random assignment of teachers or classes if the same classes serve alternately as experimental and control groups. For example, one teacher might teach his or her class using an experimental treatment for eight weeks, while a colleague with a similar class might teach using a control method. After eight weeks, the two teachers could test their students, and then switch methods. If learning is higher in the experimental condition both times the classes are compared (controlling for pretests or other covariates), we can be confident that the treatments made the difference (not just differences between teachers or classes per se).

This list is not intended to be exhaustive, but is meant to illustrate that field experiments high in external validity and low in bias can be conducted in school settings on a relatively small scale.

Practitioner-Originated Research

Any of the experimental designs described above could be implemented by the teachers and principal of a single school, or a few schools working together, with help from the district for data analysis (although with the advent of microcomputers in school buildings, teachers might be able to do this as well). Component-building research would tend to address the issues on which principals and teachers make decisions every year: Should they have a "homework every day" policy? Should they use films more frequently? Should they use a reward-for-effort or continuous progress grading system? Should they regroup for math? These are all practical questions of instructional design that would have application to all schools. Instead of having every school re-invent the wheel, schools that engaged in component-building research could communicate their findings in journals, at conventions, and so on, so that others could learn from their experiences.

Regardless of who does them, component-building studies on practical issues of instructional design could make a substantial contribution to research-based school improvement. When a large number of such studies are done, we will understand how variations in instructional components influence student achievement and other outcomes. We might finally have a science of instructional design, based on rigorous experimentation in schools, so that it will be possible to directly apply our knowledge to teacher training for instruction in settings like those in which the experiments were done. The most important practical outcome of such a program of research might be to enable us to construct complete programs far more effective than those available today, because each component of these programs will be known to increase student achievement. Some day, teachers and administrators may have a wide variety of well thought out, thoroughly evaluated instructional methods known to accelerate student achievement. To reach this goal will take extraordinary programmatic efforts on the part of many researchers, teachers, and administrators, and a consistent focus on instructional improvement on the part of funding agencies (especially NIE). It will not be quick or easy. But if educational research is justified on the grounds that it leads to improvements in practice, how can we do otherwise?

References

Anderson, L.M., Evertson, C., & Brophy, J. An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 1979, 79, 193-223.

Becker, W. & Carnine, D. Direct instruction: An effective approach for educational intervention with the disadvantaged and low performers. In B. Lahey & A. Kazdin (eds.), Advances in Child Clinical Psychology. New York: Plenum, 1980.

Block, J.H., & Anderson, L.W. Mastery Learning in Classroom Instruction. New York: Macmillan, 1975.

Brophy, J.E. Teacher behavior and its effects. Journal of Educational Psychology, 1979, 71, 733-750.

Burstein, L. The analysis of multi-level data in educational research and evaluation. In D. Berliner (ed.), Review of Research in Education (vol. 8). Washington, D.C.: American Educational Research Association, 1980.

Campbell, D., & Stanley, J. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Carroll, J.B. A model for school learning. Teachers College Record, 1963, 64, 723-733.

Denham, C., & Lieberman, A. (eds.) Time to Learn. Washington, D.C.: National Institute of Education, 1980.

Ebeling, B. A meta-analysis of programmed instruction in grades 7-12. Paper presented at the Annual Convention of the American Educational Research Association, 1981.

Frederick, W. and Walberg, H. Learning as a function of time. Journal of Educational Research, 1980, 73, 183-194.

Glaser, R. Teaching Machines and Programmed Learning II: Data and Directions. Washington, D.C.: National Education Association, 1965.

Good, T., & Grouws, D. Teaching effects: A process-product study in fourth grade mathematics classes. Journal of Teacher Education, 1977, 28, 49-54.

Good, T., & Grouws, D. The Missouri Mathematics Effectiveness Project: An experimental study in fourth grade classrooms. Journal of Educational Psychology, 1979, 71, 355-362.

Jamison, D., Suppes, P., & Wells, S. The effectiveness of alternative instructional media: A survey. Review of Educational Research, 1974, 44, 1-67.

Karweit, N. A reanalysis of the effect of quantity of schooling on achievement. Sociology of Education. 1976, 49, 236-246.

Karweit, N. The organization of time in schools: Time scales and learning. Paper presented at the Conference on School Productivity and Effects, San Diego, 1978.

Karweit, N. Time on task: A review of the literature. Paper prepared for the National Commission for Educational Excellence, National Institute of Education, 1982.

Karweit, N., & Slavin, R.E. Measurement and modeling choices in studies of time and learning. American Educational Research Journal, 1981, 18, 157-171.

Kepler, K. and Randall, J. Individualization: Subversion of elementary schooling. The Elementary School Journal, 1977, 70, 358-363.

Kerlinger, F. Behavioral Research: A Conceptual Approach. New York: Holt, Rinehart, & Winston, 1979.

Kounin, J. Discipline and group management in classrooms. New York: Holt, Rinehart, & Winston, 1970.

Lahaderne, H. Attitudinal and intellectual correlates of attention: A study of four sixth-grade classrooms. Journal of Educational Psychology, 1968, 59, 320-324.

Marliave, R., Fisher, C., & Dishaw, M. Academic learning time and student achievement in the B-C period. Far West Laboratory for Educational Research and Development, 1978. Technical note v-29.

Miller, R.L. Individualized instruction in mathematics: A review of research.

Mathematics Teacher, 1976, 69, 345-351.

Rosenshine, B. The master teacher and the master developer. Paper presented at the annual convention of the American Educational Research Association, New York, 1982. (a)

Rosenshine, B. The Exemplary Center for Reading Instruction (ECRI) reading program. Paper presented at the annual convention of the American Educational Research Association, New York, 1982. (b)

Slavin, R.E. Separating incentives, feedback, and evaluation: Toward a more effective classroom system. Educational Psychologist, 1978, 13, 97-100.

Slavin, R.E. Cooperative learning. Review of Educational Research, 1980, 50, 315-342.

Slavin, R.E. Research Methods in Education: A Practical Guide. Englewood Cliffs, NJ: Prentice-Hall, in press.

Slavin, R.E., Leavey, M., & Madden, N.A. Combining student teams and individualized instruction in mathematics: An extended evaluation. Paper presented at the Annual Convention of the American Educational Research Association, Montreal, 1983.

Slavin, R.E., Leavey, M., & Madden, N.A. Combining cooperative learning and individualized instruction: Effects on student mathematics achievement, attitudes, and behaviors. Elementary School Journal, in press.

Stallings, J.A., & Kaskowitz, D. Follow-Through Classroom Observation Evaluation, 1972-1975. Menlo Park, CA: Stanford Research Institute, 1974.

Thompson, R. Individualized reading: A summary of research. Educational Leadership, 1975, 33, 57-61.

Wang, M. Adaptive instruction and maximizing the effective management and use of school time for instruction and learning. Paper presented at the Annual Convention of the American Educational Research Association, New York, 1982.

Zoll, E. Research in programmed instruction in mathematics. The Mathematics Teacher 1969, 62, 103-110.