WESLEY C. BECKER University of Oregon

Teaching Reading and Language to the Disadvantaged—What We Have Learned from Research

In late 1967, Project Follow Through was reorganized to select, test, and evaluate promising but different educational programs for disadvantaged youngsters in the first three grades. Now, nearly ten years later the completed evaluations of Follow Through suggest that one of these programs, the University of Oregon's Direct Instruction Model, has produced significant gains in measures of positive affect, basic skills, and conceptual reasoning. In this article, Wesley Becker discusses the distinctive features of this model—its underlying assumptions and basic teaching components. He then explores the implications of teaching reading and language skills to economically disadvantaged children and advocates that immediate steps be taken to teach vocabulary systematically throughout the school years. Viewing this goal as essential for compensatory education, he concludes with an analysis of how vocabulary instruction might best be implemented.

The teaching of reading and language competencies is at the heart of the educational phase of the War on Poverty begun in 1964. A basic assumption of this unprecedented social-action program was that the American dream of equal opportunity could be made a reality. Until recently, the evidence to support this assumption has been noticeably lacking.

As early as 1968, Daniel P. Moynihan appeared before the House Subcommittee 011 Education to summarize the 1967 Title I evaluation by the TEMPO Division of the General Electric Company:

> We had thought (as legislation such as Title I was passed) we knew all that really needed to be known about education in terms of public support, or at the very least, that we knew enough to legislate and appropriate with a high degree of confidence. . . . We knew what we wanted to do in education, and we were enormously confident that what we wanted to do could work. That confidence . . . has eroded. . . . We have learned that things are far more complicated than we thought. Thereafter simple input-output relations which naively, no doubt, but honestly, we had assumed to obtain in education simply, on examination, did not hold up. They are not there. (Cited in McLaughlin, 1975, p. 49)

The results of subsequent Title I evaluations (see, for example, Gamel, Tallmadge, Wood, & Blinkley, 1975; Glass, 1970; U.S. Office of Education, 1969, 1976) did little to challenge the view Moynihan expressed. In some measure Title I did succeed in equalizing educational opportunity by improving financial and educational resources. Nonetheless, the general failure of its programs to show consistent,

---

Reprinted from *Havrrd Educational Review, Vol. 47, No. 4, November 1977*

replicable improvements in basic and cognitive skills left educational reformers in a quandary. To show any positive effects, projects had to be selectively chosen and then superficially examined the next year (Hawkridge, Chalupsky, & Roberts, 1968).

Field-based experiments can fail for many reasons, as some have suggested. Senator Robert Kennedy suspected, for example, that the early failures of Title I were due to the disinterest if not malfeasance of local school administrators (McLaughlin, 1975). Alice Rivlin thought it was due to the technical inexperience of local evaluators (McLaughlin, 1975). From our present perspective, several reasons can be suggested why such projects seem to fail: some programs are not well designed initially (Engelmann, 1975); the self-protective bureaucratic structure of school systems resists change (McLaughlin, 1975); the instruments used to measure outcomes reveal changes caused by maturation rather than school instruction (Becker & Engelmann, 1976); some control groups simply make use of children left over after the disadvantaged are placed in special programs and, in doing so, fail to randomly select and assign students (Haney, Note 1). The lesson is clear: when field programs fail, especially if program installation and operation are not monitored, we have only a heyday for speculation. We learn little of value about those details needed to design better schools.

By late 1967, the consistent failure of outcome research to find any positive program effects was already evident, leading members of the Office of Education to take a more analytic approach in designing Project Follow Through. Efforts were first made to initiate Follow Through as a program designed to extend Head Start into the elementary grades, but in the wake of a major funding cutback, from 120 million to 15 million dollars, Follow Through shifted its aim from service to research. With this shift in funds, Follow Through was deliberately organized to select, test, and evaluate promising, but different, educational programs for disadvantaged youngsters in the first three grades. This approach was later called a planned variation design. The individual programs were installed and monitored by their originators or sponsors. Communities electing to participate in Follow Through made choices among the different models and then worked with the developers to implement specific programs. Contractual arrangements were established to encourage community implementation, and independent evaluators were given contracts to monitor the programs and evaluate the outcomes (Rivlin & Timpane, 1975). In 1968 John Hughes, then director of the Division of Compensatory Education, felt this decision marked a significant departure from earlier practices:

> The decision to use the very limited funds available for Follow Through . . . to initiate a program which will permit examination in depth of the consequence of different program approaches holds promise of inaugurating what could be literally a new era in government support for educational and social ventures, *i.e.*, an era in which the knowledge and technical expertise of the educational specialist, the systems engineer, and the behavioral scientist are brought into harmony with the pluralistic value structure of our society. (Cited by Richard Elmore in Rivlin 8c Timpane,1975, p.23)

The first director of Follow Through, Robert Egbert (Note 2), later remarked:

> "With such limited funds it seemed sensible to change Follow Through's primary purpose from 'service to children' to 'finding out what works'. . . . Follow Through now focused its attention on developing, examining and refining alternative approaches to the education and development of young disadvantaged children." (pp. 4-5)

Follow Through eventually came to involve 180 communities, 7,000 children a year, twenty-two sponsors, and an annual budget of fifty-nine million dollars. Garry McDaniels (1975), who designed the final Follow Through evaluation plan, called the program the largest and most expensive social experiment ever launched. He described the sponsors in these words:

> The sponsors represented a range of opinion, theory and rhetoric. Bank Street College of Education in New York City came with a long history of child development philosophy, theory, and practice. Siegfried Engelmann came with his learning theory and experience in highly engineered materials and teacher behaviors. Ira Gordon came with his commitments to parent training as the major vehicle for assisting children. Leonard Sealey brought "open education"; David Weikart brought a cognitively oriented curriculum. (p. 5)

Now, nine years later, the completed evaluations of Follow Through indicate that some educational programs can make a difference in teaching reading, math, and language skills. One of the models, the University of Oregon Direct Instruction Model, shows significant program gains in measures of basic skills, cognitive skills, and positive affect (Abt Associates, 1976, 1977). Given these outcomes, we can now identify more positively the problems that have been hindering Title I and other programs for the economically disadvantaged. As stated previously, when programs fail, we usually learn little about the details that can make a difference; however, with success in specific areas using certain methods, we should be able to make more precise inferences about the critical features of effective programs. The rest of this paper, therefore, will explore the implications of the Direct Instruction Model for teaching reading and language skills to economically disadvantaged children. We will first describe the model briefly.

The Direct Instruction Model [1]

The major goal of the Direct Instruction Model is to improve the basic education of children from economically disadvantaged backgrounds and thus increase their life options. Developed by Engelmann and Becker, the model had its roots in Bereiter and Engelmann's experimental preschool and in Becker's behavioral research on classroom management. The model emphasizes small-group, face-to face instruction by a teacher using carefully sequenced, daily lessons in reading, arithmetic, and language.[2] These lessons utilize modern learning principles and advanced programming strategies (Becker, Engelmann, & Thomas, 1975a, 1975b). Each set of lessons has been meticulously field-tested to determine that low-performing children will achieve the program objectives under carefully monitored conditions.

Four assumptions underlie the model. First, all children can be taught, regardless of their developmental readiness or background. Teaching failure is not excused. Second, learning the basic skills, including logical procedures, is central to intelligent behavior and should be essential to any compensatory education program. Third, disadvantaged youngsters tend to be behind other students in skills needed to succeed in school as they are now structured. Fourth, in order to "catch up," the disadvantaged must be taught more in the time available than advantaged children. The second and third assumptions appear to support a deficit view of poor children. Relative to an educational system established and controlled by the middle class, this view appears to be true; if it were not, there would have been no need for Follow Through. It should be strongly asserted, however, that children from economically disadvantaged backgrounds have a host of functional skills that are adaptive for the settings in which they live and that they are very teachable children. As environmentalists, we recognize the relativistic aspects of culture and realize that what is reinforced and functional in one culture or subculture may not be in another. We view Follow Through as an educational intervention for three or four years that can have the effect, on the average, of increasing the achievement of the economically disadvantaged and thus increasing their life options as well.[3]

The model contains seven essential teaching components. First, the teaching of general cases is emphasized so that children can generalize to all members of a set after being introduced to only a few members of that set. Second, since people are the primary instruments of instruction, the number of classroom instructors is increased. By adding teaching aides, more instruction can be given, especially to non-readers. Third, the daily program is carefully structured; when time is allocated according to teaching priorities, everyone knows what to do and when to do it. Fourth, rapid-paced, teacher-directed, small-group instruction is employed as an efficient way to individualize instruction for the non-reader. Fifth, positive approaches are used to secure and maintain student attention, reinforce correct responses, and identify mistakes. Sixth, teaching staffs are carefully trained and supervised to ensure that appropriate skills have been provided and are maintained. Seventh, student progress is monitored by means of biweekly criterion referenced tests and reports-of-lessons-taught, both of which help to detect problems while there is time to correct them.

Nine curricular strands make up the DISTAR model. Collectively they specify the teaching objectives for students in reading, arithmetic, and language. In Reading I and II, reading is taught directly with the emphasis first on decoding skills and then comprehension. In Reading III, the children are taught how to learn from reading and to extract and use new information in stories that primarily have a scientific orientation. These stories make rules available that can be used to solve problems in such areas as astronomy, muscle function, and measurement. The student completing Reading III is prepared to learn from upper-level textbooks as long as the new vocabulary and concepts in those texts are taught in some reasonable way.

Arithmetic I teaches basic addition and subtraction first through a problem solving approach; next, to prepare for more elaborate problems, children speed up adding and subtracting by memorizing number facts. In Arithmetic II, the students are introduced to multiplication and fractions; addition and subtraction skills are extended, and a variety of measurement concepts involving time, money, length, and weight are taught. The students are also taught how to work story problems, and how to derive unknown facts from those already known. In Arithmetic II, the students are taught algebra, factoring, and division in addition to traditional computation.

Language I and II teach names, classes, properties, and relational terms. Children are taught to make complete statements, to ask questions in order to explore unknown subjects, and to logically describe the world around them. These language programs emphasize comprehension and language production, and students are taught such elements of logic as conditionality, causality, multiple attributes, definitions, deductions, synonyms, and opposites. Language III expands the logical use of language and teaches basic grammatical rules. In addition, many activities in Reading and Language are also geared to building writing and spelling skills.

In developing the model, concern was expressed that children learn arts, crafts, social skills, and values in ways designed to suit local conditions. It was further stressed that instructional methods lead to a sense of personal competence and a positive attitude toward self. A positive self-concept was viewed as a by-product of good teaching rather than as a goal that could be achieved in the abstract. The National Evaluation data support this assumption (Abt Associates, 1977).

Some Distinctive Features of the Model

Many of the distinctive features of the Direct Instruction Model have been openly criticized. Thus, in addition to reviewing some of these features, this section will examine common criticisms of the model and explore appropriate counter-arguments. These issues have, however, been discussed in more depth elsewhere (Engelmann & Becker, Note 3).

Scripted Presentation of Lessons

The use of explicitly detailed lessons—scripts—has been criticized as restricting teachers' initiative. This may be a valid criticism, but one should consider the potential advantages of scripts in providing quality control in a delivery system. The scripts permit the selection and testing of sequences of examples that produce efficient learning if followed. Most teachers simply do not have time to find appropriate words and examples or to sequence skill hierarchies in the most efficient possible manner. When teachers phrase their own questions, they may choose terms unknown to lower-performing children or may include unnecessary verbiage. In choosing examples, moreover, they may teach incorrect rules because the positive examples have some irrelevant feature in common. In sequencing, it is easy to omit those skills critical for later, more complex tasks.

Another advantage of scripts is their potential for teaching teachers about effective classroom instruction. Teachers can learn effective presentation strategies through repeated examples, and it is not uncommon to find teachers using Direct Instruction techniques in subjects where scripts are not available. A most critical advantage, moreover, emerges during training and supervision. The precise skills needed to teach particular kinds of lessons can be specified when designing training programs. A supervisor entering a classroom can quickly determine what is happening and compare this with what should be occurring. The supervisor, therefore, is better equipped to provide direct, practical demonstrations or suggestions to the teacher or aide. By standardizing the teaching program in this way, it is also easier to monitor the progress of the children with criterion-referenced tests that children should pass if they have completed lessons at a specified level.

Small-Group Instruction

The use of rapid-paced, teacher-directed, small-group instruction has often been criticized as pushing or placing too much on young children. The data on affective outcomes, however, do not support this conclusion (Abt Associates, 1977). The use of small groups has many advantages. It is more efficient than one-to-one instruction and provides better teacher-direction, supervision, and individualization than large group instruction. It also emphasizes oral communication skills, which children from non-English and economically disadvantaged backgrounds often need. Finally, small-group instruction provides a setting in which the repetitious practice necessary for some important skills is made more fun by transforming drill into a challenging game. Children enjoy the rapid pacing when circumstances allow them to be successful.

Signals

Prepared scripts direct teachers how and when to use signals to cue a group to respond together. For example, in sounding out a word, a finger is used to point to the letter being sounded out. The children say the sound as long as the teacher touches it. The teacher moves his or her finger from sound to sound as they are to be said and lifts the finger away at the end of the word. To coordinate counting-to-a number, a foot tap or finger snap might be used as the signal. Often the signal simply involves dropping the hand which was raised to alert the children to "get ready" when a question was asked. Much training is required to learn how to use signals in a natural and clear manner.

The use of signals and choral response, however, has been criticized for fostering an authoritarian role for the teacher. The logic of this interpretation is questionable: carried to a proper conclusion, it would further imply that using signals to direct such groups as choirs or orchestras also promotes submissiveness or authoritarian models. An examination of children's behavior in our classrooms does not support this interpretation. Rather, the evidence tends to suggest that without signals some children merely imitate a "leader" rather than learn the task directly. The use of signals obviates this problem.

The use of signals, finally, has also been criticized as fostering rote learning. The persons drawing such conclusions, however, often overlook the long-term benefits of carefully sequenced teaching examples. The data on cognitive outcomes do not support thc criticism of rote learning (Abt Associates, 1977).

Reinforcement

Some educators believe that children should learn for intrinsic rewards. When this does not happen, however, teachers need to know HOW to use methods that will lead to learning for its own sake. The model encourages the use of positive reinforcement as a means of strengthening motivation and making learning a positive activity in its own right. When children do not respond to games, praise, attention, or success, stronger reinforcers may be necessary. Point systems leading to special rewards are sometimes needed early in a program and are faded out when no longer needed. Principles and procedures which have been the basis for teacher training in this area are now published (Becker, Engelmann, & Thomas, 1971, 1975a).

Training and Supervision

A primary focus of training in this model, therefore, is developing the skills required to teach small groups. To get the most out of each child, teachers also learn procedures for grouping the children and for changing groups. A one-week preservice workshop and continuing inservice training for one or two hours a week is usually adequate for beginners. Manuals to be followed by trainers and participants "program" the training, just as the scripts help to "program" classroom instruction (Becker, Note 4). The most common training procedure follows three steps: *model*, *lead*, and *test*. Teachers are first shown how a particular task is taught; next, they teach the task with the trainer; and, finally, they practice the task on their own in pairs, each member taking turns being student and teacher. Skilled teachers provide classroom supervision of apprentice teachers and emphasize positive reinforcement.

Biweekly Reports

Attendance information is provided biweekly, as well as results on continuous progress tests on lessons covered in reading, language, and arithmetic. These reports help teachers detect problems they themselves may have in some skill areas, or problems individual children may have. Copies of biweekly reports are further used by teachers, supervisors, and project directors to make changes in student groupings and to plan inservice training. Copies also go to the sponsor for use in monitoring site progress and for process research studies. These reports gauge the quantity and quality of the children's progress.

The Direct Instruction Model: Evaluation Results

The twenty communities which have used the Direct Instruction Model include a cross section of lower socioeconomic groups: rural and inner-city Blacks, rural whites, Mexican-Americans, Spanish-Americans, Native Americans, and students from a variety of ethnically mixed communities. At any given time approximately 8,000 low-income students were in the program. Test results from students who entered the program in thirteen sites from kindergartens in 1969, 1970, and 1971 are given in Title 1. Students were included in the analysis if they met Office of Education poverty guidelines, if they started the program at its earliest grade level, and if tests were available at more than one point in time.[4]

The Tests

Three tests were used. The first, the Wide Range Achievement Test (WRAT), measures reading, arithmetic, and spelling achievement. The reading test is quite reliable and has been demonstrated to reflect instruction that teaches decoding skills. The arithmetic test has questionable reliability and validity for some levels but provides a gross measure of computational skills. The spelling test has reasonable reliability and validity. The second test, the Metropolitan Achievement Test (MAT), assesses reading comprehension, word knowledge, math computation, math concepts, math problem solving, spelling, and language (usage, punctuation, and sentence-type). The Metropolitan Achievement Tests have excellent reliability and adequate validity, and provide measures of some of our program's goals. The third test, the Slosson Intelligence Test (SIT), is a short, individually administered test that aims to measure aptitudes similar to those measured by Stanford-Binet. We included it to measure attainment of some of the more general program goals such as language use and reasoning ability. The WRAT and the SIT have been given to nearly all students from the end of the first project year. The MAT was added in the spring of 1972 and was used each year at the end of the first, second and third grades.

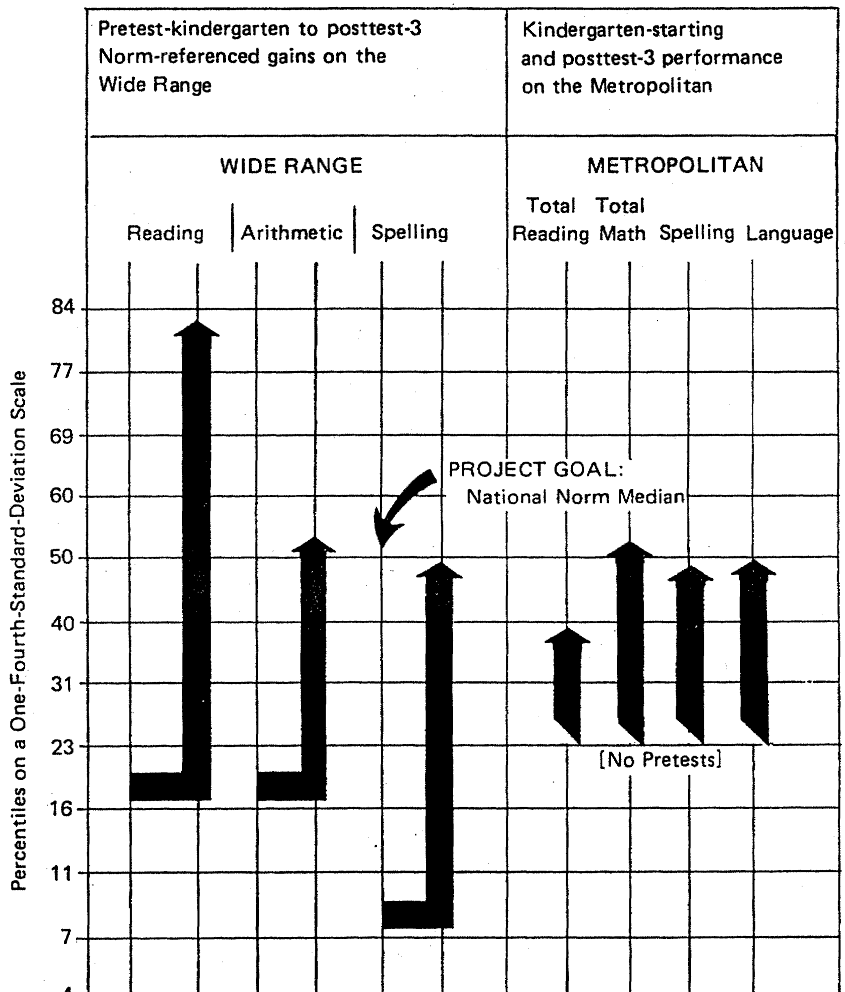
The Results

A major goal of the Direct Instruction Model has been to teach skills that would place Follow Through students above or competitive with national norms by the end of third grade. Table I shows the extent to which low-income students who started the program in kindergarten achieved this goal. This table displays what is called a "norm-referenced comparison" (Horst, Tallmadge & Wood, 1975). The intervention group is compared to the test constructor's norm group at pretest and posttest. This procedure assumes that, without the intervention program, relative positions would remain the same, on the average, and that posttest scores could be predicted from pretest scores. We assume, further, that a gain in percentiles then shows that the program is working better for the intervention group than the hypothesized average program used for similar-performing students in the normative sample.

Note that we have scaled the graph in one-fourth standard deviation units but have used percentiles to describe the scale points. This provides the reader unfamiliar with standard deviations a ready reference to percentiles but at the same time shows the magnitude of the gain in statistically more relevant units. In the National Evaluation of Follow Through, Abt Associates (1976, 1977) adopted a convention of judging an effect *educationally important* if the intervention group exceeded the comparison group by one-fourth standard deviation. At the end of third grade, such an effect would mean an approximate gain of three- to four-tenths of a grade level for most measures of achievement. We have adopted a similar convention in our own analysis because with large samples

Table I

Direct Instruction Model Follow Through Students



	Pre K	Post 3	Pre K	Post 3	Pre K	Post 3	Post 3	Post 3	Post 3	Post 3
Number of Children	2435	1988	2676	1989	2520	1974	1863	1804	1824	1854
Percentiles	18	83	19	54	8	49	40	52	48	50
Mean Norm-Group Based Standard Scores [SS]	85.9	114.5	86.6	101.2	79.5	99.7	58.0	71.1	62.8	69.9
Standard Deviation of SS	16.2	20.9	15.8	8.6	20.4	13.9	10.3	11.9	11.3	13.2
Mean Grade Equivalents	.18	5.35	.19	3.99	−.09	3.79	3.31	3.98	3.68	4.28

Descriptive statistics were computed with standard scores from tables provided by the test publishers. The mean standard scores were then converted to percentiles using the publisher's tables.

very small differences can be statistically significant but of no practical value.

For all measures on the WRAT, discernible pretest-to-posttest gains in percentiles are present. Overall in reading, low-income students moved from the 18th percentile on the pretest to the 83rd percentile on the posttest. As a group, low-income students are more than one year above the national norm in grade equivalents on these reading skills by the end of third grade. On the arithmetic portion of the WRAT, the group advanced from the l9th to the 54th percentile. In spelling, we found a gain from the 8th percentile to the 49th percentile. On these last two measures, low-income students have caught up with the national norm. The children who have been in the program only three years instead of four, perform on the average one-quarter standard deviation lower.

We have no pretest for the Metropolitan Achievement Test measures. By using the WRAT data, however, and comparable Title I data on low-income students, we have estimated entry-performance levels of our children to average no higher than the twenty-fifth percentile. We have used a slashed-arrow "pretest" point to illustrate this projection in Table I.[5] At the end of third grade, low-income Direct Instruction students performed at or near the national norm on all measures. On the Metropolitan Total Reading, a test which measures reading comprehension, students fell 10 percentile points short of the mean.

The data in Table I were collected by our research staff at the thirteen sites where children began the program in kindergarten. The National Evaluation of Follow Through by Abt Associates (1976, 1977), based on studies of seven of these sites that started Follow Through in 1970 and 1971, produces nearly identical data on the Metropolitan Achievement Test. When the site medians are averaged, using an equal weighting for each site, the Abt data show Total Reading at the 41st percentile, Total Math at the 48th, Spelling at the 51st, and Language at the 50th (Abt Associates, 1976, 1977; Becker & Engelmann, in press).

The National Evaluation also shows that, when compared to control groups, the Direct Instruction Model produces more statistically and educationally significant differences on tests of basic skills, cognitive-conceptual skills, and affective measures than any other of the eight major models.[6] Thus, in terms of level of achievement and in terms of comparisons with control groups, the Direct Instruction Model was found to be effective. After pointing out that the reading section of the Metropolitan Achievement Test is not as amenable to successful intervention as the math section, the Abt Associates evaluation (1977) concludes:

> Direct Instruction, Behavior Analysis (University of Kansas) and Bank Street models produce predominantly non-negative effects, that is, progress in reading which is either greater than or equal to the progress of comparison children. Only the children associated with the Direct Instruction Model appear to perform above the expectation determined by the progress of the non-Follow Through children. Moreover, the Direct Instruction children are the only group which appears to make more progress in reading, both early and late. In general, most models appear to be more effective during kindergarten and first grade than during second and third grade. The Direct Instruction Model is the only program which consistently produces substantial progress. (pp. 154-155)

To provide a better context for understanding this success, the performance levels of the Direct Instruction Model of the two closest programs and the average of eight major models, excluding Direct Instruction, are shown in Table II.

These positive findings imply that the Direct Instruction Model has solved some of the problems of delivering effective programs to school systems.

Implications for Improving Schools

These recent data confirm the original assumption of the advocates of the War on Poverty: poor children can succeed in school when better teaching methods are used. The question is, what are the critical elements of these methods? The seven components of the Direct Instruction Model outlined earlier concentrate on changing administrator and supervisor behavior, teacher behavior, and academic programs. We will discuss each of these areas in turn. When considering curricular suggestions, bear in mind that we are talking about teaching elementary school children with a high proportion of low-income parents. The suggestions we make, however, are not restricted to one model or only one way of doing things but focus on strategies that could work with several different methods of implementation.

Administrator and Supervisor Behavior

Tucker (Note 6) designated Behavior Analysis and the Direct Instruction as the two models that were the most successful overall. Each program channeled the actions of supervisors into helping teachers attain specified academic goals. Each program, further, nurtured the concerns of building-level administrators about their children's progress, provided feedback on that progress, and equipped supervisors with the technical skills needed to help teachers foster learning. In view of these findings, product-oriented management and better teacher-support systems appear to be essential for improving the education of the disadvantaged.

Teacher Behavior

Behavior Analysis and the Direct Instruction models carefully detail the components of effective teacher performance and trained teachers in the prescribed procedures. Training manuals were prepared for preservice and

Table II

Percentiles for Equally Weighted Site Medians for Major Sponsors Included in the National Evaluation of Follow Through

Metropolitan Test	Direct Instruction	Behavior Analysis	Bank Street	Average of Eight Sponsors (excluding DI)
Total Reading	41	34	30	24
Total Math	48	28	19	16
Spelling	51	49	32	28
Language	50	22	23	20

Data computed by the author from tables for each sponsor on results for cohort ll (Abt Associates, 1976) and cohort lll (Abt Associates, 1977) for kindergarten-starting sites. When an averaging of raw scores is weighted according to sample size, percentiles differing by a few points are obtained; but there are no remarkable changes in the ordering of outcomes.

inservice use that outlined practices for efficient management of time and specified teaching methods to be used in various situations.

Recent classroom research, relating teacher behavior to student outcomes in reading and math, provides additional support for the use of highly specified behavioral goals in the Direct Instruction Model (Brophy & Evertson, 1974, 1976; Clark, Gage, Marx, Peterson, Stayrock, & Wynne, 1976; Soar, 1973; Stallings, 1975; Stallings & Kaskowitz, 1974). These studies show that more is learned in a given subject both when more time is devoted to teaching that subject area *and* when the teaching procedures are effective. Rosenshine (1976) interprets these recent studies as suggesting that a "direct instruction model" is the most effective approach to teaching basic skills (p. 364). Direct Instruction, as defined by Rosenshine, involves teacher-directed oral activities in small groups centered on specific teacher questions, a high rate of student response, and adequate teacher feedback to students.

In a subsequent paper, Rosenshine and Berliner (in press) focus on the critical importance of "academic-engaged time" as it affects the relationship between teaching practices and outcome scores in reading and math. Correlations on the order of .50 are found between academic-engaged time and achievement-test results. Rosenshine (Note 7), in reanalyzing data collected by Stallings and Kaskowitz (1974) on six Follow Through models, found a strong relationship between the amount of engaged time a program provided and achievement outcomes. Sponsored programs that did poorly on achievement measures did not allocate as much time to teaching reading and math. Rosenshine also found that only the Oregon and Kansas models showed more academic-engaged time than the non-Follow Through control groups. A simple, logical conclusion follows: the learning of reading and math does not occur unless instructional time is provided, and the students are engaged during that time.

Academic Programs: Teaching the General Case

The Abt IV Report on Follow Through (1977) shows that in terms of grade-level scores and comparison to control groups, the Direct Instruction Model was the only major model effective on measures of academic cognitive skills (MAT Reading, MAT Math Problem Solving, and MAT Math Concepts). The University of Georgia's Mathemagenic Activities Model also did well on these cognitive measures, but since it was used in only three sites it was not considered a major model. The Kansas Behavior Analysis Model was comparable on basic skills (decoding, spelling, math computation) but not on cognitive skills.[7] We believe these findings are a function of the programming strategies used by our group but not by Kansas or the other models.

We believe that principles underlying DISTAR offer an important basis for an improved theory of programmed instruction. These principles were first discussed by Engelmann in *Conceptual Learning* (1969) and considered in more detail in *Teaching 2: Cognitive Learning and Instruction* (Becker, Engelmann, and Thomas, 1975b). In common with other theories of programmed instruction, Engelmann's specifies teaching one thing at a time, providing adequate practice, and designing lessons for a low error rate. Error rate is kept low by analyzing where errors are likely to occur, by using task-design procedures which reduce possible errors, and by pilot testing (Becker *et al.*, 1975b). Engelmann's most salient contributions to programming theory are in specifying procedures needed to teach general cases and to cumulatively build knowledge within sets of related concepts; here errors are more likely to be made because shared properties make discriminations more difficult.

There are three major kinds of general cases: concepts, operations, and problem solving rules. A general case has been taught when, after learning some members of a set, the student can generalize to all members (cf. Wittgenstein, 1958, paragraph 208).[8] In his analysis of teaching concepts, Engelmann begins with the premise that the programmer must define a concept within a specified universe of concepts: a concept is what is uniquely common to the stimulus properties of a set of instances within a specified universe of concept instances (Engelmann, 1969). Thus, the definition of a concept may change as the universe or set of concepts changes. If the only examples in a universe are *horses* or *boxes*, the learner has many options for choosing which details discriminate *horses* from *boxes*. The concept *horse* in this universe is described by the entire set of differences between *horse* and *box*. When we add *dogs* to the universe, the concept *horse* changes because many of the details that would allow the learner to discriminate between *horse* and *box* will not allow him to discriminate between *horse* and *dog*. The concept *horse* changes again when we add ponies to a set that now consists of *horses*, *boxes*, *dogs*, and *ponies*.

The notion of a concept that changes as the composition of a universe changes leads to the principle that, in order to make teaching simpler, the programmer must control the universe within which a concept "grows." For example, the student's initial concept of *vehicles* might be "something that can take you places." If instances and noninstances consist only of man-made things, this definition is adequate. When *horses* and *elephants* enter the set, the definition is revised to "*manmade* things that can take you places." Engelmann's definition of a concept makes the programming of instruction largely a logical process, requiring analysis of the distinctive features of finite sets of related concepts.

Given this logical framework, concept learning is viewed as a multiple-discrimination problem rather than a generalization problem: discriminating *between* instances and non-instances on relevant characteristics of each and discriminating *within* instances or noninstances on relevant and irrelevant characteristics. Note that Piaget (1956) refers to relevant characteristics of concepts as invariants and irrelevant characteristics as variants. Key programming principles help determine which discriminations are most important to teach and how to sequence examples for teaching them (Becker *et al.*, 1975a, 1975b).

Concepts are the important building blocks for intelligent behavior on the stimulus side. Overt operations—operant responses, or what Piaget calls schemas—are the key building blocks on the response side. Operations are defined by the common effects of a set of responses. For example, no two lifting responses are the same, but there is a common effect produced by lifting responses: an object is moved in an upward direction. The mental operation of adding involves references to possible instances of the overt operation of joining two or more groups to make one. An important step in the analysis of concepts and operations for programming instruction is the identification of building blocks. Building blocks are identified as the smallest set of stimulus/response (concept/operation) units that can be recombined to provide the largest number of applications. In oral language, a most useful building block is the referent word unit, and the general case is the sentence or utterance of two or more words with meaning. Grammatical rules enter as a second set of building blocks in constructing the general cases we call sentences. In decoding regular-sound words, a good building block is the symbol/sound unit, and the general case is the orally read word. In writing, a good building block initially is the letter-name/written-letter unit and the general case the written word. With consolidation of learning, the word-name/written-word may become the unit and written sentences the general case.

Problem-solving rules consist of sequences of operations that can be used to solve problems of a particular type. After having been taught some element of the problem set, students should be able to do any element. Assuming, for example, that the concepts *equal* and *plus* have been taught, as well as symbol identification and making lines for numbers, addition problems of the kind shown below can be taught using Rules 1-4.

2 + 4 = □
|| ||||

Rule 1. Find the side to start on. (That side cannot have an unknown.)

Rule 2. Make lines for each number of that side.

Rule 3. Count the lines.

Rule 4. Make the sides equal by placing the number counted in the box on the other side.

By teaching one additional skill, counting from-a-number-to-a-number, the students can also do problems of the form shown in the equation below.

2 + □ = 6
|| ||||

Rule 1. Find the side to start on.

Rule 2. End up with the same number on the other side by saying the first number "Two" and counting to 6, making a line under the box for each count.

Rule 3. Count the lines under the box and write the number in the box.

Similarly, in teaching decoding in reading, one can teach a set of sounds, blending skills, and rapid pronunciation skills, so that the student can read any regular sound word composed from the sounds taught (Carnine, 1977). This method of teaching decoding involves a problem-solving rule chain and yields a great saving in teaching time.

These examples display some of the analytic strategies Engelmann uses in building programs to teach the general case. His work represents a fusion of cognitive and behavioral theory. Cognitive theorists will recognize that Engelmann's approach encompasses cognitive processes. The instructional program, however, focuses on the observable behavior with which the teacher and programmer must deal in order to teach children (see Becker *et al.*, 1975a, 1975b; Becker & Engelmann, Note 3; Engelmann & Carnine, Note 8).

Academic Programs: A Problem Linear Set

Probably the most important implication to arise from our nine years of work in teaching more than 25,000 disadvantaged students is that schools are designed to teach middle-class children and need to be redesigned for teaching all students. This observation bears out our initial assumptions. However, in our preoccupation with effective instruction from kindergarten through the third grade, we failed, until now, to explore the broader ramifications of this assumption. In the analysis that follows, we conclude that schools systematically fail to provide instruction in the building blocks crucial to intelligent functioning, namely, words and their referents. Children from homes where there is strong adult support for refining the use of language are more likely to succeed in school than those from homes with less adult-child contact and adults with less education (Coleman, 1975; Freeberg & Payne, 1967; Glass, 1973).

We came to this conclusion when trying to understand our failures and successes. As we have argued, there is little to be learned from programs that fail. When instruction is effective in some areas, however, more precise inferences are possible. Consider the pattern of findings in Table I. Children in the Direct Instruction Model score at the 83rd percentile on the WRAT Reading; this is a full standard deviation above the national median. But on the MAT Total Reading these same children score at the 40th percentile, a quarter-standard deviation below the national median. In arithmetic, spelling, and language they score at the median. How are these differential findings to be explained?

WRAT Reading is a measure of decoding skills. If decoding is taught by teaching forty sounds and by teaching how to blend sounds rapidly together, a general case can be learned that will permit the reading of any regular-sound English word. A great efficiency in teaching is achieved when a relatively small set of building blocks can be recombined into a large set of applications. The DISTAR method proceeds with these assumptions and is remarkably effective in teaching decoding, especially in comparison to sight-word programs (which were most probably used for a majority of the WRAT norm group).

The levels of achievement in spelling, math, and language can be viewed as products of general cases with smaller ranges of application. The ability to reach the national average in spelling is most likely an outcome of teaching reading-by-sounds. The related process, spell-by-sounds, is an effective procedure for spelling words with regular letter-sound correspondences; the process also assists with irregular words. In our model, arithmetic is taught as several problem-solving skills that operate as general cases. These cases have more limited applications than decoding, however, and thus generate less net gain. The language skills, involving general cases of grammatical rules and word classifications, appear to have about the same level of potency as arithmetic.

The MAT Total Reading is made up of two subtests at the elementary level: Word Knowledge and Reading. The former is a test of vocabulary, the latter a test of comprehension. One is not likely to comprehend texts without knowledge of word meanings. The vocabulary-concept load[9] of the MAT Reading Test (Elementary Level, Form F) is beyond the experience of most disadvantaged third graders. Consider such items as *country* when used to mean nation, *Amazon ant*, *probably*, *exterminator*, *penicillin*, *disease-causing germs*, *Egyptians of old*, *a seated cat*. The test also demands that students make logical inferences and deductions.

In contrast to the general-case learning involved in decoding, arithmetic, grammar, and spelling-by-sounds, the learning of vocabulary and concepts usually involves a "linear-additive set" (Becker & Engelmann, 1976, p. 58). In a linear additive set, the learning of one element gives little advantage in learning a new element. To be sure, there are families of words that have common root meanings and common meanings of affixes, which permit some limited general cases to be generated. But, by and large, the learning of proper names, new concepts, and the learning of synonyms for concepts already known by another name, involve linear additive sets in which each new element must be taught. Knowledge of the English language, which is absolutely essential to oral and written comprehension, serves largely to define intelligent behavior (Miner, 1957). Teaching this language involves a task of the first magnitude.

The magnitude of the problem is revealed by the following analysis. Basal reading texts control vocabulary instruction for three years. Chall (1967) indicates that about 1,500 words are covered in this time, although the average student may actually have an oral vocabulary that is two to three times this size. After the third grade, schoolbooks shift to an uncontrolled adult vocabulary. Thorndike and Lorge (1944) estimate that the average high-school senior knows about 15,000 words, including proper nouns and derivatives but not inflections. Dupuy (1974) estimates, however, that the average high-school senior knows about 7,000 basic words. Basic words do not include proper nouns, derivatives, inflections, and compounds. On the average, 5,000 of these basic words are learned after third grade. Since schools do not systematically build vocabulary knowledge, students with weak home training in English language skills are often in trouble by fourth grade, if not earlier.

In a recent review of research on socioeconomic factors in intelligence, Terhune (1974) details the correlates of IQ and school achievement (and by implication language comprehension). General language competence is low when: a single parent is in the home; per-capita income is low; education of caretakers is low; birth order of subject is high; and the number of children is large. If we add to this list, "parent's primary language is not English or is not standard English," a set of conditions is described that occurs all too frequently among those called "disadvantaged." The hypothesis that vocabulary-concept knowledge plays a major role in reading comprehension is supported in the research literature. In his review of research Carroll (Note 9) concludes that

> much of the failure of individuals to understand speech or writing beyond an elementary level is due to deficiency in vocabulary knowledge. It is not merely the knowledge of single words and their meanings that is important, but also the knowledge of the multiple meanings of words and their grammatical functions. (p. 175)

Carroll also argues that vocabulary-concept knowledge is the key area of concern for improving reading comprehension for the economically disadvantaged, but he is quite aware that there is no easy way to make gains in this area.

Chall (1958) noted much earlier that "of the diverse stylistic elements that have been reliably measured and found significantly related to difficulty, only four types can be distinguished: vocabulary load, sentence structure, idea density, and human interest." (p. 157) Of these factors Chall suggests that vocabulary load "is most significantly related to all criteria of difficulty so far used." (p. 157)

Other analyses of our Follow Through data tend to support the view that vocabulary-concept knowledge is not systematically taught by schools. Analysis of the MAT Reading tests indicates a progressive loss on percentiles from end of grade one to end of grade three, which is paralleled by a progressive change in the tests toward an adult vocabulary by the end of third grade. On MAT Total Reading, low-income children in our program are far above the median at the end of first grade (70th percentile), but they drop progressively by the end of second grade (56th percentile) and the end of third grade (40th percentile). This drop is found for all reading subtests—Word Knowledge, Word Analysis,[10] and Reading (Becker & Engelmann, Note 5). Such losses are not found in decoding, math measures, or spelling, in all of which substantial percentile gains occur during kindergarten and first grade are then maintained.

One might argue that the basis for these differential effects lies in the logical requirements of reading comprehension. This argument, however, does not explain why the effect is also found for Word Knowledge but is not found for Math Problem Solving and Math Concepts. From the point of view of programmed instruction, the central problem is the linear-additive set, vocabulary. Logical processes, such as those involved in some aspects of comprehension and math, are general case sets; these can be taught with some efficiency, but vocabulary does not submit to such methods.

When the children taught by the Direct Instruction method are followed up in fifth and sixth grades, they maintain superior performance levels on decoding and spelling but lose about one-half standard deviation against the norm group on MAT Total Reading and Total Math (Becker & Engelmann, Note 5). The losses on reading and math are consistent with the hypothesis that fourth-, fifth-, and sixth-grade teachers are not building on the skills the children had when they left third grade. However, effective programs in math are available to schools; they just are not being used or used effectively.

Schools have never had programs in reading that systematically build vocabulary concept knowledge. Except for technical vocabularies, this task has been largely left to parents. Furthermore, since the achievement tests are built by procedures that measure age changes, and not simply the effects of school instruction, children from homes with weak support for language development fall progressively behind on current reading tests. This finding is commonly reported. When student gains on such instruments are used to evaluate teachers, moreover, those teaching children with higher socioeconomic status (SES) will appear more competent than those teaching lower SES children. The tests are not built to be specifically sensitive to school instruction but to age changes generally (Becker & Engelmann, 1976).

IQ and learning gains of the children in the Direct Instruction Model were analyzed according to IQ blocks (i.e., 71-90, 91-100, 101-110). The data show that children with high entry IQ's start higher on most achievement measures and stay higher but do not usually gain proportionately more (Becker & Engelmann, Note 5). Exceptions to this pattern are the first-year gains on WRAT Reading (decoding) and the gains from the end of second to the end of third grade on MAT Total Reading (comprehension). In both instances higher IQ groups show proportionally greater gains than lower IQ groups. This second exception is centrally important. If we assume that the IQ test is measuring general language competence as developed at home and school, children with greater language competency should excel on the adult-level vocabulary required by the MAT Reading Test at the elementary level. Since a more controlled vocabulary is used for the first- and second-grade tests, language competency, as measured by general IQ, should not provide much advantage on these tests if there is effective school instruction.

These logical and empirical analyses clearly point to a problem for educators who strive to teach reading comprehension to all children. The data suggest that school programs do not systematically build vocabulary-concept knowledge. Current programs are structured to teach middle-class children or children who, to a large extent, are taught oral-language comprehension at home. We assume that this form of language learning is then transferred to reading comprehension at school.

Implications for Curriculum Design

Redesigning curricula to solve the problem of vocabulary-concept learning will not come easily. The first goal is to define and analyze a reduced lexicon in a way that will facilitate its teaching. Part of the analytic task is restructuring what has been treated as a linear-additive set into general-case sets. Miller and Johnson-Laird (1976) argue that such restructuring is a difficult but possible task: "We assume that at the heart of the sprawling and ever changing English lexicon some nuclear, widely shared set of words and concepts provides a framework for the development of special vocabularies for special cases." (p. 668) Thus, from the point of view of programming, the task is to specify a minimum, adult-level competency and then systematically build to that goal. The extensive work of Thorndike and Lorge (1944) on word frequencies, the analyses by Ogden (1932) on basic English, the work by Dupuy (1974) on basic words, as well as linguistic research on roots and affixes, all contribute information that could be used to build a system for teaching a basic lexicon or, more simply, a basic vocabulary.

In particular, Dupuy's contribution has been to provide a definition of a basic word which excludes proper nouns, derivatives, inflections, compounds, archaic words, foreign words, and technical terms. Within this framework, he has demonstrated that there must be about 12,300 basic words in English since he found 123 basic words in a 1 percent sample of the 240,000 words in *Webster's Unabridged Dictionary*. He has further shown that the average high-school senior knows about 7,000 of these words (70 words in his initial sample multiplied by 100).

It would, of course, be a major undertaking to repeat Dupuy's analysis 100 times to find these 7,000 basic words. It seems very likely, however, that the basic words known by the average high school senior would be contained in the 30,000 words identified by Thorndike and Lorge (1944) as being most frequent in printed English.[11] Thus, by analyzing the Thorndike-Lorge list for basic words, researchers could considerably short-cut the search. Furthermore, an analysis of this kind would produce a list of high-frequency proper nouns to be included in a vocabulary program and indicate what savings might potentially be made in teaching derivatives, inflections, and compounds. As one indication of how this analysis might proceed, consider this example: The Thorndike-Lorge list contains the words *negate, negative, negation*, and *negatively*. *Negate* could be treated as a basic word and probably could be taught through carefully presented examples. This decision, however, would depend on what other word-concepts have already been taught. The other words could then be treated as related vocabulary words and taught through general cases dealing with suffixes (*-ive, -ion,* and *-ly*). Of course, such teaching would have to provide adequate practice within diverse contexts.

Given a vocabulary divided into basic and related words, two types of analyses would be necessary before one could start to engineer a program to efficiently teach vocabulary. First, on the *form side*, the problem would be to identify the morphemes (the smallest units of meaning) contained in the 7,000 basic words. If the basic words could be shown to be made up of only 2,000 to 2,500 morphemes, a potential savings in the analysis and teaching of word *forms* would be possible. Dixon's *Morphographic Spelling Program* (1976) is based on this potential. By teaching 720 morphographs, Dixon has been able to generate spellings for more than 10,000 English words using only five rules for combining morphographs.

Consider, for example, the following sequence, by frequency, of words from the Thorndike-Lorge list: *help, support, insist, September, toil, resist, recognizable, assistance. September* would go into a special list of proper nouns. The others could, in part, be analyzed as shown in Example I.

Many more related words could be found in the Thorndike-Lorge list that would improve the ratio of basic words to related words, and similarly, other basic words, such as *detract*, could introduce morphemes that would improve the ratio of morphemes to basic words. Further improvements in potential efficiency could result from the analysis of variants or allomorphs, such as *sup-sub, com-con,* and *im-in*. These variations can be accounted for by rules of internal morphology, including devoicing and assimilation. This suggests the possibility of teaching such variations as general cases (Dixon & Becker, Note 10).

The second type of analysis needed to teach vocabulary efficiently would involve the semantic or *concept side* of the word-concept field. The morphographic analysis discussed above could easily provide a basis for determining which words are semantically the sum of their parts and which are not. For example, the components of *compress* quite literally mean "push together"; collectively they provide a direct meaning for compress. The components of *sarcastic*, on the other hand, mean "related to one who is of the flesh" and thus give little due to the word's meaning.

The analysis of morphographic meaning would assuredly aid in making some program decisions, but a more general approach to the "semantic core" of the lexicon will be clearly needed if the overall task is to become feasible. What is needed is a core vocabulary-concept set that can be initially taught through examples and then used as definitions to teach other word meanings. Ogden's (1932) basic English provides one such core. The system consists of 850 words and guidelines for combining them into meaningful, standard-English sentences. Although Ogden devised the system as a basis for scientific and commercial communication between the nations of the world, it has great potential for use in developing programs to teach English as a primary or secondary language. The translation of *Robinson Crusoe* and the Bible into basic English gives evidence of the system's ability to deal with the conceptual complexities of the language.

It seems reasonable to propose, therefore, that each of Dupuy's estimated 12,300 basic words be defined in Ogden's basic English. For example, *to*, *give*, *help* and *work* are considered by Ogden to be basic English words. These may be used to define *assist* as "to give help," or *toil* as "to work." Through computer analysis, groupings of words with related basic definitions could then be identified to provide a basis for developing efficient procedures to teach the semantic side of the language. The vocabulary program envisioned here would first teach the concepts contained in basic English, or test for their knowledge, and then build systematically on this base. The goal, therefore, would be to develop a graded progression of vocabulary words or families of words and their major meanings, where, as much as possible, a word introduced at a given level can be defined in terms of words already introduced. In addition, a systematic plan for introducing inflections, derivatives, compounds, proper nouns, and idioms would have been designed, and cross-checked against other major analyses of reading vocabularies (see, for example, Harris & Jacobson, 1972).

Given Dupuy's findings that the average student at the end of third grade knows about 2,000 basic words, and the average twelfth grader knows about 7,000 basic words, the system would have to teach 550 basic words a

Example 1

Basic	Class	Related	Class	Morphemes
help	v,n	helper, etc.	n	help + er
support	v,n	supporter	n	sup + port + er
insist	v	insistence	n	in + sist + ence
toil	v,n	toiler	n	toil + er
resist	v	resistance	n	re + sist + ance
recognize	v	recognizable	nm	re + cogn + ize + able
assist	v	assistance	n	as + sist + ance

(v=verb, n=noun, nm=noun modifier)

year, or twenty-five per school week. The addition of derivatives, compounds, and proper nouns would probably double this requirement but not double the required teaching time. When we recognize that all new learning must be carefully discriminated from prior, related learning, such programs pose no small teaching task. The problem is further complicated by the range of individual differences within any grade between four and twelve. By grade twelve, this range approaches 5,000 basic words (Dupuy, 1974). By the use of carefully structured programs to boost vocabulary competency for low-performing children in the early grades, the number of children in the lower end of this range can be reduced. By structuring school programs to teach basic operations in the various areas of knowledge using basic words, the advanced children would not necessarily be held back.

Given that a basic vocabulary can be defined and graded using empirical and logical criteria, a "free-use" vocabulary could be further defined and used as a guide for preparing texts for upper-grade levels. Free-use words would be those that should be known by at least 80 to 90 percent of the students at a particular grade. Words in a proposed text that are not suitable for a given grade level would be replaced, emphasized in the text, or listed so that the teacher could teach them before beginning a lesson. Now that books can be printed through the use of magnetic tape, a computer program could be used at the level of galley proofs to permit changes or at least provide a supplemental vocabulary listing. The goal would be to furnish text systems that progressively build a knowledge of basic vocabulary throughout the school years. With statements of goals and "free-use rules," all current texts could immediately be analyzed by the author and publisher to provide a vocabulary-building supplement that any teacher could use. Finally, the analysis could be used to produce vocabulary-concept programs by grade levels that could be components in the teaching of reading comprehension. This could be especially important for those children whose language training at home has been insufficient. Such programs, moreover, would also have considerable use in preparing handicapped children for mainstreaming and could take a variety of forms. Perhaps the lower-level programs would resemble thc current Direct Instruction programs, whereas middle-level programs would utilize more self-instructional devices such as programmed reading or language masters, and advanced programs would be taught within content areas but coordinated across areas and levels.

Conclusions

A technology exists to achieve the educational goals of the War on Poverty. The Direct Instruction Model has demonstrated that children from low-income homes can be taught at a rate sufficient to bring them up on most achievement measures to national norms by the end of third grade. The model has been most effective in specifying objectives that can be taught as general-case strategies for decoding in reading, solving math problems, and making logical inferences. The model has also been more effective in expanding vocabulary knowledge and other comprehension skills than other models studied by Follow Through. The model, however, did fall somewhat short of its goals to reach national norms in the comprehension area. Our analysis leads to the conclusion that educators have failed to systematically teach words and their referents—among the most important building blocks for intelligent functioning. The DISTAR language and reading programs provide a basis for improvement in this area but do not go far enough. Children coming from home backgrounds that fail to provide adequate training for the continuous growth of vocabulary and concepts are likely to continue to fall behind in public schools. Even after four years of intensive effort, children taught in the Direct Instruction Model are far short of an average high-school senior's vocabulary—a vocabulary that is presumed by most fourth-grade tests. The data on our fifth and sixth graders strongly imply that the schools fail to build skills toward this goal. One might conclude that this is simply more evidence of the failure of compensatory education. When viewed from a different perspective, however, the data point to a failure of school programs to accomplish their assigned roles.

Advocates of compensatory education assumed that all learning problems would be solved by finding the critical stage where some magic could be applied to fertilize cognition. The Follow Through data suggest that a magical solution is unlikely. Massive restructuring of school systems is required. As schools are presently constituted, there is no way that low-performing children can be adequately prepared for the vocabulary they will encounter by the fourth grade. Clearly, the first step toward improving this situation is to recognize that language learning does not end by the third grade. Once this is understood, programs can then be engineered to teach vocabulary-concept knowledge in a systematic way throughout the school years. These programs can also accelerate the learning of logical operations through general-case teaching. Research has shown the job can be done with preschool children, with children in kindergarten through third grade, and with older children in remedial programs. If approached systematically, this job can be done for all.

Reference Notes

1. Haney, W. *Analysis of previous Follow Through national evaluation reports* (Draft Report). Cambridge, Mass.: The Huron Institute, 1976.
2. Egbert, R. *Planned variation in Follow Through*. Paper presented to The Brookings Institution Panel on Social Experimentation, Washington, D.C., April 1973.
3. Engelmann, S., & Becker, W. C. *Forward from Basics-the case for direct instruction*. In preparation.
4. Becker, W. C. *Direct instruction implementation manuals* (Follow Through Project). Unpublished manuscript, University of Oregon, 1977.
5. Becker, W. C., & Engelmann, S. *Analysis of achievement data on six cohorts of low income children from 20 school districts in the University of Oregon Direct Instruction Follow Through Model* (Follow Through Project, Technical Report #76–1). Unpublished manuscript, University of Oregon, 1976.
6. Tucker, E. *The Follow Through planned variation experiment*. Paper presented at the meeting of the American Educational Research Association, New York, April 1977.
7. Rosenshine, B. *Academic engaged time, content covered and direct instruction*. Paper presented to the American Educational Research Association, New York, April 1977.
8. Engelmann, S., & Carnine, D. W. *Analysis and psychology of instructional design* (tentative title). In preparation.
9. Carroll, J. B. *Learning from verbal discourse in educational media: A review of the literature*. Princeton, N.J.: Educational Testing Service, 1971. (ETS RM 71–61).
10. Dixon, R. & Becker, W. C. *A proposal for the analysis of a basic vocabulary to facilitate instruction in the language arts*. Unpublished manuscript, University of Oregon, 1977.

References

Abt Associates. *Education as experimentation: A planned variation model* (Vol. III). Cambridge, Mass.: Abt Associates, 1976.

Abt Associates. *Education as experimentation: A planned variation model* (Vol. IV). Cambridge, Mass.: Abt Associates, 1977.

Becker, W. C., & Engelmann, S. The direct instruction model. In R. Rhine (Ed.), *Encouraging Change in America's schools: A decade of experimentation*. New York: Academic Press, in press.

Becker, W. C., & Engelmann, S. *Teaching 3: Evaluation of instruction*. Chicago: Science Research Associates, 1976.

Becker, W. C., Engelmann, S., & Thomas, D. R. *Teaching-a course in applied psychology*. Chicago: Science Research Associates, 1971.

Becker, W. C., Engelmann, S., & Thomas, D. R. *Teaching 1: Classroom management*. Chicago: Science Research Associates, 1975. (a)

Becker, W. C., Engelmann, S., & Thomas, D. R. *Teaching 2: Cognitive learning and instruction*. Chicago: Science Research Associates, 1975. (b)

Brophy, J. E., & Evertson, C. N. *Process-product correlations in the Texas teacher effectiveness study. Final report*.

Austin, Texas: University of Texas, 1974. (ERIC Document Reproduction Service No. ED 013 371)

Brophy, J. E., & Evertson, C. N. *Learning from teaching*. Boston: Allyn & Bacon, 1976.

Carnine, D. W. Phonics versus look-say: transfer to new words. *Reading Teacher*, 1977, 30, 636–640.

Chall, J. S. Readability: An appraisal of research and application. Bureau of educational Research Monograph 1958, 34. Columbus, Ohio: Ohio State University.

Chall, J. S. *Learning to read: The great debate*. New York: McGraw-Hill, 1967.

Clark, C. M., Gage, N. L., Marx, R. W., Peterson, P. L., Stayrook, N. G., & Winne, P. H. *A factorially designed experiment on teacher structuring, soliciting, and reacting. Final Report*. Stanford, Calif.: Center for Research and Development in Teaching,1976. (ERIC Document Reproduction Service No. ED 134 591)

Coleman, J. S. Methods and results in the IEA studies of effects of school on learning. *Review of Educational Research*, 1975, 45, 335-386.

Comptroller General Report to the Congress. *Assessment of reading activities funded under the federal program of aid for educationally deprived children*. Washington, D.C.: U.S. Office of Education, December, 1975.

Dixon, R. *Morphographic Spelling Program.* Eugene, Ore.: E-B Press, 1976.

Dupuy, H. F. *The rationale, development, and standardization of a basic word vocabulary test*. Washington, D.C.: U.S. Government Printing Office, 1974. (DHEW Publication No. [HRA] 74-1334)

Engelmann, S. *Conceptual learning*. San Rafael, Calif.: Dimensions Publishing, 1969.

Engelmann, S. *Your child can succeed*. New York: Simon and Schuster, 1975.

Freeberg, N. E., & Payne, E. Parental influence on cognitive development in early childhood: A review. *Child Development*, 1967, 38, 65-87.

Gamel, N. N., Tallmadge, G. K., Wood, C. T., & Blinkley, J. L. *State ESEA Title I Reports: Review and analysis of past reports, and development of a model reporting system and format* (Report UR294). Bethesda, MD.: Research Management Corporation, 1975. (ERIC Document Reproduction Service No. ED 118 684)

Glass, G. V. Problems in implementing the Stull Act. In N. L. Gage (Ed.), *Mandated evaluation of education: A conference on California's Stull Act*. Stanford, Calif.: Stanford Center for Research and Development in Teaching, 1973. (Distributed by Capital Publications, Educational Resources Division, Washington, D.C., 1974)

Glass, G. V., & others. *Data analysis of the 1968-1969 survey of compensatory education (Title I)*. Washington, D.C.: U.S. Office of Education, Bureau of Elementary and Secondary Education, August 1970. (ERIC Document Reproduction Service No; ED 057 146)

Harris, A. J., & Jacobson, M. D. *Basic elementary reading vocabularies*. New York: Macmillan, 1972.

Havighurst, R. J. Poverty sub-cultures-The faces of rural poverty. *In Education for the culturally disadvantaged* (Proceedings of the National Conference on educational objectives for the culturally disadvantaged). Hot Springs, Arkansas: South Central Region Educational Laboratory, September 1967. (ERIC Document Reproduction Service No. ED 024 735)

Hawkridge, D. G., Chalupsky, A. B., & Roberts, A. O. H. *A study of selected exemplary programs for the education of disadvantaged children* (Parts. 1 and 2). Palo Alto, Calif.: American Institute for Research in the Behavioral Sciences, September 1968.

Horst, D. P., Tallmadge, G. K., & Wood, C. T. *A practical guide to measuring project impact on student achievement* (Monograph No. 1 on Evaluation in Education). Washington, D.C.: U.S. Government Printing Office, 1975.

McDaniels, G. Follow Through yields new evidence of school effects. *Education Daily*, March 20, 1975, p. 5.

McLaughlin, M. W. *Evaluation and reform*. Cambridge, Mass.: Ballinger, 1975.

Miller, G. A., & Johnson-Laird, P. N. *Language and perception*. Cambridge, Mass.: Belknap Press of Harvard University Press, 1976.

Miner, J. B. *Intelligence in the United States*. New York: Springer, 1957.

Ogden, C. K. *The basic words: A detailed account of uses*. London: Kegan Paul, 1932.

Piaget, J., & Inhelder, B. The child's conception of space. London: Routledge & Kegan Paul, 1956.

Rivlin, A. M., & Timpane, P. M. *Planned variation in education*. Washington, D.C.: The Brookings Institution, 1975.

Rosenshine, B. Classroom instruction. In N. L. Gage (Ed.), *The psychology of teaching methods* (75th NSSE Yearbook). Chicago: University of Chicago Press, 1976.

Rosenshine, B. V., & Berliner, D.C. Academic engaged time. *British Journal of Teacher Education,* in press.

Soar, R. S. *Follow Through classroom process measurement and pupil growth (1970-71). Final report.* Gainesville, Fla.: College of Education, University of Florida, 1973. (ERIC Document Reproduction Service No. ED 113 288)

Stallings, J. A. Implementation and child effects of teaching practices in Follow Through classrooms. *Monographs of the Society for Research in Child Development,* 1975, 40, (7-8, Serial No. 163).

Stallings, J. A., & Kaskowitz, D. H. *Follow Through classroom observation evaluation (1972-73).* Menlo Park, Calif.: Stanford Research Institute, 1974.
Terhune, K. W. A review of the actual and expected consequences of family size (Colspan Report No. DP-5333-G-1, 1974). Washington, D.C.: U.S. Government Printing Office, 1974.
Thorndike, E:. L., Lorge, I. *The teacher's word book of 30,000 words.* New York: Bureau of Publications, Teachers College, Columbia University, 1944.
United States Office of Education. *Education of the disadvantaged: An evaluation report on Title I, Elementary and Secondary Education Act of 1965, Fiscal Year 1968.* Washington, D.C.: U.S. Office of Education, Bureau of Elementary and Secondary Education, 1969
United States Office of Education. *Annual evaluation report on programs administered by the U.S. Office of Education, FY 1975.* Washington, D.C.: Capital Publications, Educational Resources Division, 1976.
Wittgenstein, L. P*hilosophical investigations* (2nd ed.). New York: Macmillan, 1958.

1 This summary is adapted from Becker & Engelmann (Note 3).
2. The programs are published by Science Research Associates, Chicago, under the trade name DISTAR© and were designed by Siegfried Engelmann.
3. We do not assume minority children have poor language skills. We do assume that there is a higher probability of weaker language skills in children in lower socioeconomic settings (and some of the data presented later support the assumption). In our Follow Through data, this weakness is found least in the big cities of the Northeast (see Havighurt, 1967), but there compensatory language programs have been considered needed and helpful.
4. This data base provides the largest sample size for measurement of program impact. To check for biases in the outcomes because of attrition, we have also analyzed year-to-year gains and pre-to-post gains (K to 3 or 1 to 3) on the same students. These additional analyses do not materially change any conclusions except to make the actual gains somewhat greater than those reported here. When the percent of the children who did not meet OEO poverty guidelines are added to the analysis, there is another slight increment in the level of performance. Children who enter the program late perform a year lower on the average. This would be expected if the program is effective. (For more details on sponsor data see Becker & Engelmann, Note 5).
5. A U.S. Office of Education report (1976) substantiates this approach: "Analyses of all test scores showed that the typical student who received compensatory assistance in reading was at the 20th percentile for grade 2 and the 22nd percentile for grades 4 and 6." (p. 88) Moreover, in a footnote to page 88 this additional information is given: "In conjunction with the Emergency School Aid Act evaluation, children in grades 3, 4, and 5 of a nationally representative sample of minority isolated schools (50% or more non-white) performed at the 23rd, 18th and l9th percentiles, respectively, On reading achievement in the Spring 1973; similar results were obtained for mathematics achievement" (Ozenne, D. G., *et al.,* 1974). The educational requirement for Title I eligibility (one year or more below grade level is thc 20th percentile for Metropolitan Total Reading.
6. In the National Evaluation report Abt Associates (1977), Basic Skills arc defined by scores on the Metropolitan subtests for Spelling, Math Computation, Word Knowledge, and Language(grammar). Cognitive Skills are defined by scores on MAT Reading, Math Concepts, Math Problem Solving, and Raven's Coloured Progressive Matrices. Affective Measures arc defined by scores on thc Coopersmith Self Esteem Inventory and the Intellectual Achievement Responsibility Scale(IARS). Thc Coppersmith is designed to assess children's feelings about themselves, about what others think of them, and about School. The IARS measures the extent to which children attribute their success or failures to themselves or to outside forces.
7. This conclusion is based in part on data from thc Wide Range Achievement Test used in early grades by thc National Evaluation and by sponsors. Decoding, however, is not measured by the tests used for Abt Associates (1977).
8. Note also that the term "general case" covers the linguist's term "generative." Behavioral procedures have been used to experimentally produce "generative" verbal behavior in a number of studies (see Becker *et al.,* 1975h, 85-86).
9. We use the term "vocabulary-concept load" to make clear that we are referring not just to verbal behavior but also to knowledge of referents.
10. This subtest is found only on the first- and second-grade tests.
11. The Thorndike-Lorge listing gives an overall estimate of word frequencies in written materials, based on averaging frequencies for four samples: (1) the Thorndike general count of 1931, (2) the Lorge magazine count, (3) the Thorndike count of juvenile books, and (4) the Lorge-Thorndike semantic count.