DOCUMENT RESUME

ED 288 921 UD 025 894

AUTHOR	Madden, Nancy A.; Slavin, Robert E.
TITLE	Effective Pull-out Programs for Students at Risk.
INSTITUTION	Center for Research on Elementary and Middle Schools, Baltimore, MD.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Apr 87
GRANT	OERI-G-86-0006
NOTE	34p.; Paper presented at the Annual Meeting of the American Educational Research Association (April 20-24, 1987, Washington, DC).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; Class Size; *Compensatory Education; Computer Assisted Instruction; Elementary Education; *High Risk Students; *Individualized Instruction; *Instructional Design; *Program Design; Program Effectiveness; Remedial Instruction
IDENTIFIERS	Education Consolidation Improvement Act Chapter 1

ABSTRACT

This paper reviews research on effective pull-out programs for elementary students who are at risk for school failure. The emphasis is on programs that can be readily replicated by schools other than those that developed them. The focus is on programs provided to students who have been identified as in need of remedial services which are implemented outside of the regular classroom. The study found that effective pull-out programs for students fall into three broad categories: (1) diagnostic-prescriptive programs; (2) tutoring programs; and (3) computer-assisted instruction. Successful examples of each of these types are described and discussed. The most successful models completely adapt instruction to students' unique needs and provide plentiful direct instruction appropriate to their levels of readiness. Results of the study suggest that the achievement of at-risk students can be significantly increased, either by making relatively inexpensive but extensive modifications in the regular instructional program or by implementing relatively expensive but intensive interventions as pull-out programs. It is possible that a combination of these strategies would be more effective than either one by itself. An extensive list of references is included, along with descriptions of successful programs among each of the three types. (PS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

11-19-87

ED288921

Effective Pull-cut Programs

For Students at Risk

Nancy A. Madden

Robert E. Slavin

Center for Research on Elementary and Middle Schools

Johns Hopkins University

Paper presented at the annual convention of the American Educational Research Association, Washington, D.C., April, 1987. This paper was written under funding from the Office of Educational Research and Improvement (Grant No. OERI-G-86-0006). However, the opinions expressed are those of the authors and do not necessarily reflect OERI positions or policy.

UD025894

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ERIC

From the beginning, one of the guiding principles of compensatory education has been that compensatory services must supplement, not supplant, the educational programs provided to eligible students. Low achieving students in schools which are eligible for Chapter I (formerly Title I) services must receive something identifiably "extra." A small army of evaluators located in every state check to see that Chapter I services are in fact supplementing rather than supplanting regular education programs.

Partly because of the "supplement, not supplant" regulations, schools have overwhelmingly relied on pull-out models as a means of providing Chapter I or Title I services. Most often, students who qualify for compensatory services are taken out of their regular classrooms for 30-40 minutes of remedial instruction in reading and/or mathematics. This arrangement has the advantage of making it clear that Chapter I services are supplementary, as special personnel and materials are clearly only allocated to identified students. In in-class alternatives to pull-out, maintaining the distinction between who is served and who is not served is more difficult. At least partly for this reason, Chapter I programs have overwhelmingly chosen pull-out as the mode of service delivery. A study conducted during the 1981-82 school year found that pull-out outnumbered in-class models by nine to one (Advanced Technology, 1983), and in a more recent study of schools specially chosen to represent a variety of service delivery models, fifteen of seventeen elementary schools used pull-out in reading and/or math (Rowan, Guthrie, Lee & Guthrie, 1986).

Despite the many criticisms of pull-out (see, for example, Glass & Smith, 1977; Johnston, Allington, & Afflerbach, 1985; Archambault, 1987), pull-out

ERIC

is likely to remain as a widely-used means of providing compensatory education services. In any case, many have noted previously (e.g., Archambault, 1987; Kennedy & Birman, 1986), the important issue is not the setting in which compensatory services are provided, but the quality of the programs provided in the setting.

This paper reviews research on effective pull-out programs for elementary students who are at risk for school failure. A companion paper by Slavin & Madden (1987) discusses effective classroom programs for the elementary grades, and a paper by Karweit (1987) discusses effective preschool and kindergarten programs.

Scope of the Review

The focus of this paper is on programs provided to students who have been identified as being in need of remedial services which are implemented outside of the regular classroom. This excludes programs implemented in the regular class (which are reviewed by Slavin & Madden, 1987) and self-contained "replacement" or special education programs in which students are assigned to a remedial class for most or all of their school day. The emphasis of this review is on programs which could be (or have been) readily replicated by schools other than those which developed them. For example, many school districts have developed exemplary Chapter I programs (see Griswo'd Cotton, & Hansen, 1986), but the features of these programs are often uniquely adapted to the situations, personnel, and students of the district and were not designed to be replicated as such in other districts. Also, a few studies have identified variables related to effective Chapter I programs (e.g., Cooley & Leinhardt, 1980; Crawford, in press). In contrast to these, the programs emphasized in this review are models which have well-specified manuals, materials, training procedures, and other features which characterize programs intended for replication by others.

Review Procedures

The procedures used in this review are described in detail by Slavin & Madden (1987). Essentially, the review procedure was an adaptation of best-evidence synthesis (Slavin, 1986), a method which combines the features of meta-analytic and traditional narrative reviews. The initial literature search examined many sources, including published and unpublished articles, school district reports, and government documents. Requests for information were sent to all of the 116 exemplary compensatory education programs identified by Griswold et al. (1986). However, the most useful sources of information were reports submitted to the Joint Dissemination Review Panel (JDRP), a U.S. Department of Education panel which reviews evaluations of programs supported by federal funds. Programs whose effects are certified by the JDRP as valid are eligible for funding and dissemination through the National Diffusion Network (NDN). The JDRP submissions are especially useful because virtually all programs submitted to JDRP are designed to be replicated by others and because the JDRP requires data that can be used to assess program effects.

A set of substantive and methodological inclusion criteria were applied in deciding which programs to emphasize in this review. These are described in detail by Slavin & Madden (1987). In brief, programs had to assess effects on standardized reading and/or math scales in studies of at least a semester's duration. Programs had to be compared to matched or randomized control groups, or year-to-year gains of at least five NCE's had to be reported (fall-to-spring NCE gains, which have been found to have serious methodological problems, were not considered adequate evidence of program effects; see Gabriel, Anderson, Benson, Gordon, Hill, Pfannenstiel, & Stonehill, 1985). All programs presented data to indicate that they were effective with students eligible for Chapter I services or other at-risk students.

Categories of Effective Pull-Out Models

The effective pull-out programs for students at risk of school failure fell into three broad categories. One, diagnostic-prescriptive programs, covers the great majority of existing Chapter I programs. In this model, students identified as being in need of remedial services are carefully assessed and then instruction appropriate to their needs is given by a teacher in a location separate from the regular classroom. Instruction may be given to individuals or to small groups within a pull-out class of roughly three to eight students.

The second category is tutoring programs, in which tutors work one-on-one with identified tutees. Tutors may be teachers, paraprofessionals, volunteers, or older students. In the third category, computer-assisted instruction (CAI), students work on computers for at least part of their remedial reading or math time. It should be noted that many diagnostic-prescriptive programs use computers for management (as opposed to instructional) purposes, but these are not categorized as CAI.

Characteristics and Outcomes of Effective Pull-Out Models

Diagnostic-Prescriptive Programs

Because of the widespread use of diagnostic-prescriptive pull-out programs, the data available on such programs is extensive. Almost all of the exemplary Chapter I programs identified by Griswold et al. (1986) used diagnostic-prescriptive models, as did many of the programs certified by the JDRP. However, very few of these presented convincing evidence of effectiveness. In most cases, these programs were identified as exemplary on the basis of fall-to-spring gains in normal curve equivalent scores (NCE's). However, most districts presented fall and spring scores for several years, and with few exceptions these indicate little or no growth in fall scores or spring scores over time. The typical pattern is a fall-to-spring gain of 8-12 points, followed by a decline over the summer of the same magnitude. This cannot be seen as convincing evidence that the programs are not effective; it could be that the lack of year-to-year growth in scores is due to dropping of students with high scores from the Chapter I lists, or that scores increased when new programs were introduced but then remained stable during the years for which data were provided, or there may be other methodological or substantive factors involved.

Table 1 Here

However, there are a few diagnostic-prescriptive programs which do present more convincing evidence of effectiveness. Information on these programs is provided in Table 1.

ERIC

Only two of the successful diagnostic-prescriptive programs used control group designs. One of these, Project Conquest (JDRP No. 74-12) mostly provides remedial services to groups of six students. However, early in the year, students receive one-to-one tutoring until they acquire word perception skills. Students are carefully assessed by special "reading clinicians," and given individual prescriptions.

An evaluation of Project Conquest was conducted in low-income, mostly black schools in East St. Louis, Illinois. Project Conquest students made greater gains than control students at grade levels from 1-6 on many standardized leading measures. However, these data are difficult to interpret because no evidence is given that the experimental and control classes were initially equivalent. A more sophisticated analysis was conducted using fifth grade data and correcting for pretest differences, and this also showed a clear advantage for the program.

One of the exemplary Chapter I programs identified by Griswold et al. (1986) used a control group design to evaluate its diagnostic-prescriptive pull-out model. This is the Oklahoma City, Oklahoma Chapter I program, a classic diagnostic-prescriptive pull-out model. In this program, students are assessed in terms of skills and learning styles and then given instruction appropriate to their needs, individually or in small groups. Some use is made of computer-assisted instruction; the evaluation of this component is discussed below in the section on CAI.

The evaluation of the Oklahoma City diagnostic-prescriptive pull-out program (Kimball, Crawford, & Raia, 1985) involved careful matching on prescores of students who received Chapter I services with those who did not.

The same procedures were followed in two successive years. Results indicated that Chapter I students gained 3.0 NCE's more in math than their counterparts who did not receive Chapter I services. Using the standard deviation of NCE's of about 12 estimated by Gabriel et al. (1985), this is equivalent to an effect size of +0.25. However, gains in reading were much smaller and, in one year, were not statistically significant (ES=+.12).

It so happens that Kimball et al. (1985) computed effects of the Oklahoma City Chapter I program for 1984-85 three different ways, using fall-to-spring, spring-to-spring, and control group designs. It is instructive to see the radical differences in the NCE gains from these three methods, summarized in the following chart:

Oklahoma City Chapter I Evaluation, Grades 1-6

NCE's Computed Three Ways

	Reading	Math
Fall-to-spring	13.03	17.05
Spring-to-spring	5.87	8.00
Experimental-control	1.90	2.80

The above chart shows clearly that different evaluation designs lead to different conclusions, and counsels extreme caution in interpreting results of fall-to-spring and spring-to-spring data.

The Baltimore, Maryland school system also conducted a comparison of students served by Chapter I pull-out programs and matched students eligible for Chapter I but not served. This evaluation found that served and unserved students made about the same fall-to-spring gains in reading and math scores (Granick, Quigley, Katzenellenbogen, & Richardson, 1984).

Other than Project Conquest, only one JDRP-approved program appears in Table 1, The Diagnostic-Prescriptive Arithmetic Program (JDRP NO. 74-68). This model, developed and evaluated in Staten Island, New York, uses a math lab approach to remediation of deficits in mathematics. Individualized and small group activities keyed to problem areas identified by the Stanford Diagnostic Mathematics Test are provided to students. Spring-to-spring gains of approximately 10.4 NCE's were made; however, since the Stanford Diagnostic Mathematics Test was also used as the posttest, there is a possibility that this program was essentially teaching the test.

Two additional programs from the list of effective Chapter I programs also presented evidence of effectiveness based on outstanding year-to-year gains. One of these is the Lincoln, Nebraska Chapter I program. In this program, a computer management system was introduced to help handle diagnostic tests, assign students to the program, maintain coordination between the regular teacher and the Chapter I teacher, monitor student progress, and evaluate student success. During the years when the program was implemented, Chapter I students showed steady gains in spring scores in reading and math which appear to be due to the addition of the computer management system to the diagnostic-prescriptive pull-out model already in use in the district (Weatherl, 1986).

Another apparently effective Chapter I model cited by Griswold et al. (1986) is the Columbia, Missouri Public Schools' Chapter I mathematics pro-

ERIC

gram. This model places considerable emphasis on coordination of instruction between Chapter I and regular classroom teachers; forms indicating specific objectives students are working on are passed back and forth between Chapter I pull-out teachers and regular classroom teachers, and time is set aside for Chapter I and regular teachers to meet. Program data indicate fall-to-fall gains of 3.7 NCE's in math.

The evidence from the Chapter I programs listed in Table 1 indicates that diagnostic-prescriptive programs <u>can</u> be effective, but their effects do not generally appear to be large. However, the almost universal use of fall-to-spring gains (or year-to-year gains at best) as the criteria of program effects means that we really know little about which programs (or components of programs) are having important effects on the students they serve.

<u>Tutoring Programs</u>.

A wide variety of tutorial programs using tutors ranging from experienced, specially trained teachers to paraprofessional aides to older, low-achieving students have demonstrated considerable effectiveness in improving students achievement in reading and math for students in grade 1 through 6. These effective programs fall essentially into two groups; those that were designed as remedial programs and those that were designed as preventative programs. Three models for remediation were identified. The two strongest models, Training for Turnabout Volunteers (TTV) and School Volunteer Development Project (SVDP), were developed in Dade County, Florida. TTV used volunteer junior high school students who took tutoring as an elective class to tutor low achieving first through sixth graders in reading and math.

Tutors were required to be able to read and compute at the fifth grade level, which allowed underachieving students to act as tutors. Tutors received considerable training involving a specified structured curriculum before they began to tutor, and spent one day out of five in group supervision, receiving continuing training in the specific subject matter being taught as well as in tutoring skills such as rewarding the tutees' successes, refraining from criticizing failures, organizing the work to be presented, and so on. The tutorial materials were not programmed. Tutors drew from a wide variety of materials deemed to be useful for teaching the needed skills. The gains made by students working with the well-trained tutors were compared to the gains made by students working with tutors who did not receive continuing supervision, but otherwise worked for similar amounts of time with similar students using the same kinds of materials. Students worked together four days a week for sixteen weeks for forty minutes a day.

Gains were significantly greater for both tutors and tutees when the tutors received continuing training. Students tutored by trained tutors in math gained .93 standard deviations more than those tutored by untrained tutors on the Metropolitan Achievement Test. In reading, the tutees of trained tutors gained .51 standard deviations more on the Metropolitan than their comparison group. Trained tutors gained .49 standard deviations more than their untrained counterparts in math, but did not gain significantly more in reading.

One difficulty with the TTV program was that it required that an elementary and junior high school be physically close to one another so that students could travel back and forth in the time allotted for one class period and still have time to tutor. The second program developed in Dade County avoided this difficulty by focusing on the recruitment and training of adult volunteers as tutors. In the School Volunteer Development Project, adult volunteers tutored underachieving first through sixth graders for a half hour a day four or five days a week. Tutors were trained prior to tutoring in a variety of tutoring skills and use of multimedia materials, and worked with the reading specialist on the skills they were tutoring. Again, a variety of materials pertaining to the skills being developed were used. Students were randomly assigned to tutored or untutored groups. Those who received tutoring gained .50 standard deviations more in reading and 1.1 standard deviations more in math than the untutored control students.

In contrast to the two programs above, Success Controlled Optimal Reading Experience (SCORE), the third successful remedial model identified, uses highly programmed materials and a very specifically structured tutoring session. SCORE uses rapid drill and practice in lists of words grouped to teach specific decoding skills (short and long vowel sounds, sound blending, word patterns such as night, light, fight, and so on). Students who are deficient in decoding skills are tutored for 15 minutes a day until they complete the program, usually a period from four to six months.

In the original evaluation of the program, educationally handicapped students were randomly assigned to a special class for learning disabled students, to regular class placement with no support, or to regular class placement with 15 minutes of tutoring by students in grades 6 to 12. The SCORE treatment group exceeded both control groups on measures of word recognition and oral reading accuracy. Effect size estimates ranged from .5

to .7 standard deviations. The results were replicated over three successive years. The program has been implemented in a wide variety of settings with lower middle and upper class populations in urban and rural settings using older students, parent volunteers and aides as tutors. Gains in these settings have exceeded the gains in the original study in almost every case.

A limitation of these data lies in the fact that no reading comprehension data are provided. The gains seen in word recognition and oral reading accuracy are large and important, and would be expected to enable readers to pay more attention to comprehending their reading, but there is no guarantee that this generalization occurs. Comprehension gains would have to be assessed in any implementation of this program. It may be that SCORE would provide a strong first step in a more comp[illegible] program that explicitly addressed comprehension as well.

Three different programs have addressed the problem of reading failure by attempting to prevent failure at the first grade level. Programmed Tutorial Reading (PTR) provides one-to-one tutoring by paraprofessionals to first graders in the bottom quartile in reading. The tutoring process is highly structured. The tutor uses programmed materials based on the basal series used in the school that instruct the tutor in where to start, what to say, when to praise, how to respond to a failure in word recognition, and so on. Students are essentially taught the words in the basal using a sight word approach, and in addition are taught word analysis and passage comprehension skills using the same vocabulary in separate series of lessons.

Programmed Tutorial Reading has been evaluated against a control group in two studies. In the first (Ellson, Harris, and Barber, 1968), students were randomly assigned to receive PTR, directed tutoring, or no tutoring. Directed tutoring involved having paraprofessionals tutor students using specific materials provided by the teacher that were designed to correspond to the instruction going on in the classroom. Tutors in both PTR and directed tutoring were trained for about 18 hours prior to the beginning tutoring. Results indicated that PTR was more effective than directed tutoring and control, but the differences were not large.

In the second study of PTR (JDRP #74-17), students recieving fifteen minutes of tutoring a day were compared to students receiving no tutoring over a one year period. In this study, substantial differences were seen between the two groups in favor of the tutored group.

Two other prevention models used programs designed to improve students' skills in specific reading or prereading areas. The Wallach Tutorial Program (Dorval, Wallach, and Wallach, 1978), used paraprofessionals to tutor children identified as low in readiness skills for one-half hour per day on phoneme identification skills. This training focused the students on breaking up the sounds in words so that, for instance, beginning sounds could be heard, sounded, and represented separately from the rest of the word. Wallach and Wallach based their program on research showing that disadvantaged children were drastically deficient in recognizing phonemes in heard words - such as knowing whether house or man starts with the sound "mmmmmmm". The tutored children performed at the 56th percentile on the California Test of Basic Skills at posttest after having begun at the 17th percentile. The control group moved from the 19th percentile to the 35th. The effect size for the comparison is .75 standard deviations.

Prevention of Learning Disabilities, a program developed by the New York University Medical Center, takes an approach that is somewhat simiair to the Wallach and Wallach approach. Students are screened for deficits in sensory skills related to reading at the kindergarten or beginning first grade level, and those found to be deficient are tutored in those areas in which they are low. Auditory discrimination exercises similar to those used by Wallach and Wallach form a significant part of this instruction. Students were assigned to work on a tutorial basis or in small groups (two to three students) with a resource teacher three to five times a week. The program is designed as a two-year program, but evaluations after both one and two years showed significant gains for the treated group on word recognition scales and on a word attack measure. No standardized tests of reading comprehension were used.

While the evidence of effectiveness for the tutoring programs is very impressive and supports the use of such models, one important piece of data is lacking. Tutoring models, in particular those using paid tutors and which identify themselves as "preventative," often justify their costs in terms of reducing the need for later remediation. However, none of the evaluations of preventative tutoring models actually present data on long-term maintenance of effects, either on achievement or on assignment to special or remedial education. One preventative tutoring program that has generated considerable interest in recent years, Reading Recovery (Clay, 1985; Boehnlein, 1987), claims to remove the need for future remediation for most of the high-risk first graders it serves, but evidence supporting this is lacking as of this writing.

Overall, the results of the tutoring studies show this strategy to have strong effects on student achievement. When older students are used as tutors, they also experience important achievement gains; when either peers or adult volunteers provide tutoring, the costs of tutoring programs are low.

Computer Assisted Instruction

Computer assisted instruction (CAI) can be seen as another form of tutoring in which the "tutor" is a machine rather than a person. Like a human tutor, computers can determine students' needs, provide instruction appropriate to those needs, recognize and reinforce student success, and keep records of student progress. However, computers are usually less able to explain concepts to students, and for this reason CAI programs invariably have teachers available to students while they are working on their computer lessons.

Several studies of CAI as a pull-out program for students in need of remediation have been conducted, but both the largest number of evaluations and the highest quality evaluations have involved the reading and math programs developed at Stanford University in the early 1970's and currently disseminated by the Computer Curriculum Corporation, or CCC (Jamison, Fletcher, Suppes, & Atkinson, 1976). Unlike most current CAI programs which operate on microcomputers, CCC uses a mainframe with terminals linked to a central processing unit by telephone. The computer keeps records of student performance levels and progress, and provides students with exercises appropriate to their needs.

The most important evaluation of the CCC curriculum itself was conducted in the Los Angeles Public Schools by the Educational Testing Service (Ragosta, 1983). In this four-year longitudinal study, students were randomly assigned to receive ten minutes per day of CAI in mathematics, reading, or language, as part of a thirty-minute Chapter I pull-out period. Results on the CTBS indicated substantial effects of CAI on math computations, which increased from an effect size of .36 in the first year to .72 by the third year of intervention. Effects on concepts and applications scales were smaller and non-significant. In reading, positive effects were found for vocabulary and comprehension scales after one year (ES=+.25 and +.23, respectively), but while the vocabulary effects increased to +.59 by the third year, comprehension effects actually favored the control group by that time (ES=-.24). Effects in language mechanics were at about one-quarter of a standard deviation all three years, but were small and generally non-significant on language expression and spelling scales.

Two JDRP-approved projects which used the CCC curriculum also present convincing evidence of effectiveness. One, a program evaluated in Lafayette Parish, Louisiana, called Title I Mathematics Laboratory with Computer Assisted Instruction (JDRP No. 82-46), randomly assigned students to regular Title I pull-out or to a combination of the regular pull-out model with ten minutes of CAI in mathematics. Effects on the CTBS Total Math Scale were modest, only +.19, but the extraordinary quality of the experiment makes the effect credible. Another JDRP program based on the CCC reading materials is the Merrimack (Massachusetts) Education Center (JDRP No. 82-34), a program which supplements ten minutes per day of CAI with tutorial and small group instruction from the teacher. A study with random assignment of students to

ERIC

CAI and control treatments found positive effects averaging 40% of a standard deviation on the MAT reading scale.

One interesting CAI program which does not use CCC materials is Basic Literacy through Microcomputers, a model developed and evaluated in Salt Lake City, Utah. This program uses either electric typewriters or computers to supplement the teacher's instruction by giving students opportunities to practice and apply phonics skills by typing words, sentences, and stories. The evaluation of this program involved comparison classes in gains on standardized tests. Differences at grades 1 and 3 favored the experimental group, but substantial pretest differences at grade 3 make these results inconclusive. Ironically, the grade 1 experiment, which found substantial positive effects on California Achievement Test Total Reading scores (ES=+.58), used typewriters rather than computers.

Overall, results for the CAI programs (especially CCC) are well-established and positive, though in the best-controlled studies they are usually modest in magnitude and appear more frequently on basic skills than on higher-order skills. Since the costs of CAI can be very high (see Ragosta, 1983), this approach can be compared to adult tutoring, which tends to have larger effects in studies of similar methodological quality.

Conclusions

In a companion paper to the present article, Slavin & Madden (1987) reviewed research on effective classroom programs for students at risk for school failure. They found that the most consistently successful classroom models were continuous-progress programs in which students are taught in skill-level groups and proceed through a hierarchical set of skills, and cooperative learning programs in which students also receive instruction at their appropriate levels but then practice skills in mixed-ability learning teams. On the basis of this and other evidence, Slavin & Madden (1987) concluded that effective programs for students at risk balanced adjustment of instructional approaches to meet students' unique needs with provision of adequate direct instruction. In addition, effective classroom programs provide frequent assessment of student progress through a well-specified, hierarchical set of skills.

The examination of effective pull-out programs conducted in the present paper provides further support for these conclusions. The most successful models, tutoring and CAI, completely adapt instruction to students' unique needs and provide plentiful direct instruction appropriate to students' levels of readiness. Diagnostic-prescriptive pull-out models, which have generally been less successful than tutoring or CAI, also carefully assess students' needs and adapt instruction to these needs, but often suffer from one of two problems.

Chapter I pull-out teachers often work with individual students, leaving others in the group to spend much time working on worksheets, which may be of relatively little value. Alternatively, Chapter I teachers may present lessons to heterogeneous groups of students which do not adapt to their individual needs. In either case, the instruction provided in traditional diagnostic-prescriptive pull-out programs may not be markedly better than that provided in the regular classroom. If so, it is unrealistic to expect that 30-45 extra minutes of instruction will make a substantial difference in achievement. Both the tutoring and the CAI studies suggest that *intensive* interventions are needed to make a substantial difference in a pull-out program.

Taken together, the conclusions of the present paper and of Slavin & Madden (1987) suggest that the achievement of at-risk students can be significantly increased, either by making relatively inexpensive but extensive modifications in the regular instructional program or by implementing relatively expensive but intensive interventions as pull-out programs. It is possible that a combination of these strategies would be more effective than either one by itself.

There is much more we need to know about effective programs for students at risk. Research on diagnostic-prescriptive models is generally of low quality; it may be that certain forms of diagnostic-prescriptive models would be highly effective in well-designed experiments, but much research and development is needed to establish what these forms might be and how much difference they could make in student achievement. The tutoring and CAI studies are of much higher methodological quality, but except in cases where peer tutoring or volunteer programs are practical, these are expensive interventions. For practical as well as theoretical reasons, then, there is a need to identify the *elements* of tutoring and CAI that account for their effects, so that perhaps these same principles could be applied in a less expensive form.

We can make an important difference in the achievement of students at risk for school failure. Of that there is no doubt. All that is in doubt is our willingness as a society to support continued research and development of effective, replicable models and to support the systematic use of the methods research has already identified.

ERIC

REFERENCES

Advanced Technology. (1983). Local operation of Title I, ESEA 1976-1982: A resource book. McLean, VA: Author.

Archambault, F.X. (1987, April). Pullout versus inclass instruction in compensatory instruction. Paper presented at the Annual Convention of the American Educational Association, Washington, DC.

Boehnlein, M. (1987). Reading intervention for high-risk first-graders. Educational Leadership, 44(6), 32-37.

Clay, M.M. (1985). The early detection of reading difficulties. Exeter, NH: Heinemann.

Dorval, B., Wallach, L., & Wallach, M.A. (1978, April). Field evaluation of a tutorial reading program emphasizing phoneme identification skills. The Reading Teacher, 784-790.

Ellson, D.G., Harris, P., & Barber, L. (1968). A field test of programmed and directed tutoring. Reading Research Quarterly, 3(3), 307-368.

Gabriel, R.M., Anderson, B.L., Benson, G., Gordon, S., Hill, R., Pfannenstiel, J., & Stonehill, R.M. (1985). Studying the sustained achievement of Chapter 1 students. Washington, DC.: U.S. Department of Education.

Glass, G.V., & Smith, M.L. (1977). Pull-out in compensatory education. Washington, DC.: Department of Health, Education, and Welfare.

Glass, G., McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Granick, L., Quigley, M.M., Katzenellenbogen, R., & Richardson, L.H. (1984). Optimizing Chapter I program effectiveness. Baltimore, MD: Baltimore City Public Schools.

Griswold, P.A., Cotton, K.J., & Hansen, J.B. (1986). Effective compensatory education sourcebook. Washington, DC: U.S. Government Printing Office.

Jamison, D., Fletcher, J.D., Suppes, P., & Atkinson, R.C. (1976). Costs and performance of computer-assisted instruction for education of disadvantaged children. In J. Froomkin, D. Jamison, & R. Radner (Eds.), Education is an industry. Cambridge, MA: Ballinger.

Johnston, P., Allington, R., & Afflerbach, P. (1985). Curriculum confluence in classroom and clinic. Elementary School Journal, 85, 465-477.

Karweit, N.L. (1987, April). Effective early childhood education and kindergarten programs for children at risk of academic failure. Paper presented at the Annual Convention of the American Educational Research Association, Washington, DC.

Kennedy, M. (1978). Findings from the Follow Through planned variation study. Educational Researcher, 7, 3-11.

Kennedy, M.M., & Birman, B.F. (1986). The effectiveness of Chapter 1 services. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Kimball, G.H., Crawford, J., & Raia, F. (1985). Evaluation of the 1984-85 Oklahoma City Public Schools Chapter I Program. Oklahoma City, OK: Oklahoma City Public Schools.

Ragosta, M. (1983). Computer-assisted instruction and compensatory education: A longitudinal analysis. Machine-Mediated Learning, 1, 97-127.

Rowan, B., Guthrie, L.F., Lee, G.V., & Guthrie, G.P. (1986). The design and implementation of Chapter 1 services: A study of 24 schools. San Francisco: Far West Laboratory for Educational Research and Development.

Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Research, 15(9), 5-11.

Slavin, R.E., & Madden, N.A. (1987, April). Effective classroom programs for students at risk. Paper presented at the annual convention of the American Educational Research Association, Washington, DC.

Weaterl, R. (1986). Lincoln Public Schools Chapter I. Lincoln, NE: Lincoln Public Schools.

Table 1: Effective Diagnostic-Prescriptive Pullout Programs

PROGRAM AND DESCRIPTION	SAMPLE	DESIGN	EFFECTS	EFFECT SIZE
<u>Programs Evaluated Using Control Groups</u>				
Project Corquest East St. Louis, 1974. (JDRP #74-12) Individual and small group instruction provided outside of class to remediate specific deficits identifed by individual assessment. Extensive inservice.	Grades 1-6, reading. Students performing below grade level.	Comparison group identified (criteria not clear); IQ's slightly lower for comparison group.	Treatment group gains in each grade exceeded comparison group gains significantly on standardized reading tests. No evidence of other adoptions.	Cannot be estimated.
Oklahoma City, OK Chapter 1 Program. (Kimball, Crawford, and Raia, 1985) District-wide management objectives used to coordinate regular class and Chapter 1 instruction.	Grades 1-8, reading and math. Chapter 1 students.	Control group selected from Chapter 1 eligible students not served.	California Achievement Test Math- trt gained 3.0 NCE's more than control Reading- trt gained 1.45 NCE's more than control	 ES=.25 ES=.12

Table 1, continued

PROGRAM AND DESCRIPTION	SAMPLE	DESIGN	EFFECTS	EFFECT SIZE
<u>Programs Evaluated Using Year-to-Year Gains</u>				
Lincoln, Nebraska Chapter I Management and Coordination Project. Developed district level objectives and a monitoring system. Computers used for management fo data. Day-to-day coordination between Chapter 1 and regular class.	Grades K-6, reading and math. Chapter 1 students.	No control group. Spring-to-spring gains measured.	Spring-to-spring gains of 4-5 NCE's in reading and [illegible] in math. System-wide year-to-year continuing gains are shown.	Not estimated.
Diagnostic-Prescriptive Arithmetic Program, 1972. Staten Island, NY. Math lab approach emphasizes hands-on experience, manipulatives.	Grades 3-5, math. Students two years below grade level. level.	No control group. Spring-to-spring gains measured (student level).	Stanford Diagnostic Arithmetic Test. Spring-to-spring gains of 10.4 NCE's.	Not estimated.
Columbia, MO. Chapter I Math Program. Emphasizes coordination between Chapter I and regular classroom, as well as use of manipulatives in math instruction.	Grades 2-6, math. Chapter I students.	No control group. Fall-to-fall scores for successive cohorts.	Iowa Test of Basic Skills. NCE gains fall-to-fall of 8.7 NCE's.	Not estimated.

ERIC

Table 2: Effective Tutoring Programs

PROGRAM AND DESCRIPTION	SAMPLE	DESIGN	EFFECTS	EFFECT SIZE
__Programs Evaluated Using Control Groups__				
__A. Remedial Programs__				
Training for Turnabout Volunteers. (JDRP #81-11) Cross-age tutoring in reading and math. Tutors trained with pre-service and inservice classes. Not programmed material. Tutored 40 min./day four days/week. In-service training on fifth day. Cost minimal - 1 (part-time) paid staff per school.	Tutors 7-9 graders Tutees 1-6 graders Five schools in Miami (Dade Cty.) Tutored in reading and math.	Compared trained vs. untrained tutors. Non-equivalent but unbiased groups. Duration of tutoring - one school year.	Tutees Metropolitan Ach. Test (Math): Tutees having trained tutors (N=62) gained significantly more than those having untrained tutors (N=62). Metropolitan Achievement Test (Reading): Tutees having trained tutors (N=51) gained significantly more than those having untrained tutors (N=51)	ES=.93 ES=.51
			Tutors MAT (Math) trained tutors > untrained MAT (Reading) trained tutors > untrained	ES=.49 ES=.08
School Volunteer Development Project (JDRP #75-79) Community Volunteer Tutors. Tutored 2-4 hrs./week/volunteer. Each student tutored 1/2 hr. 4 days/week minimum. Cost minimal - 3 paid staff per district.	Reading and math. Grades 2-6	Random assignment to tutored or untutored groups. Duration of tutoring - one school year	Metropolitan Achievement Test Reading tutored (N=119) > untutored (N=117) by 1 GE gain Math tutored (N=119) > untutored (N=117) by .85 GE gain	ES=.50 ES=1.1
Success Controlled Optimal Reading Experience (SCORE) (JDRP #80-42) San Francisco, CA Each student tutored 15 min./day by older students or adult volunteers.	Grades 1-6 Reading	Randomly assigned - 2 controls replicated over 3 yrs.	Tutored (N=61) > untutored (N=58) on word recognition (WRAT/Gilmore) NO TESTS OF COMPREHENSION	WRAT ES=.5-.7 Gilmore ES=.5

ERIC

Table 2, continued

PROGRAM AND DESCRIPTION	SAMPLE	DESIGN	EFFECTS	EFFECT SIZE
B. Preventative Programs				
Programmed Tutorial Reading (Study 1) Farmington, Utah (72-73) (JDRP #74-17) Each student tutored 15 min./day by para-professionals or older students.	Grade 1, reading, bottom quartile	Matched exp./ control students in pairs - no pretest differences Duration: 1 year	Gates McGinty Reading Scores Tutored (N=33) > untutored (N=33) Replicated using non-standardized tests in several other districts.	ES=.65 vocabulary ES=.41 comprehension ES= .50 basal comprehension measure
Programmed Tutorial Reading (Study 2) (Ellson, Harris, Barber, 1968) Each student tutored 15 min./day by para-professionals. Programmed tutoring compared to directed tutoring (comparable to the regular classroom activities done with the individual student)	Grade 1, reading, selected from entire range of first graders in the class. N=43 in each group	Sixteen schools in inner city. Random assignment to experimental groups then matched control selected from same class. Exp. trts. were: 1 or 2 sessions programmed 1 or 2 sessions directed tutoring	Programmed tutoring > directed tutoring on Stanford Ach. Test Total Reading Tutoring of any kind > no tutoring on criterion referenced measures only	ES=.10
Prevention of Learning Disabilities - New York (JDRP #79-33) Each student tutored 3-5 times/week for 30 min. by a resource teacher.	Grades 1 & 2 Students with low readiness scores.	Randomly assigned to test or control Measured at end of 1st and 2nd grade	Tutored (N=87) > untutored (N=39) WRAT Woodcock Word Attack Woodcock Word Recognition	 ES=1.00 ES=2.00 ES=1.00
Wallach Tutorial Program (Dorval, Wallach and Wallach 1978) Half-hour per day 28 hours total Tutored by para-professionals.	Grade 1 - students below 40% on MAT TRT N=20 One control group from same school (N=20) One control group from different school (N=18) Nine student control groups received extra help from an aide.	Not clear how students were assigned to treatments but groups are equivalent at pretest. Treatments were tutoring vs. no tutoring vs. time with aide.	Spache Word Rec. tutored > control Spache Passages tutored > control CTBS Total Rdg. tutored > control	Diff=1.3 GE ES=.75

ERIC

Table 3: Effective Computer-Assisted Instruction Programs

PROGRAM AND DESCRIPTION	SAMPLE	DESIGN	EFFECTS	EFFECT SIZE		
<u>Programs Evaluated Using Control Groups</u>						
Computer Curriculum Corp. Drill and practice curricula in reading, math and language arts						
Study 1: Los Angeles Unified School District 4 schools (Ragosta, 1983)	Math, Grades 1-6 Reading, 3-6 Language, 3-6	Compared to randomly selected controls. Groups studied longitudinally over 4 years.	California Test of Basic Skills (CTBS)	yr 1	yr 2	yr 3
			Math Computations	+.36	+.56	+.72
			Math Concepts	-.02	+.12	+.09
			Math Applications	+.03	+.12	+.26
			Reading Vocabulary	+.25	+.17	+.58
			Reading Comprehension:	+.23	-.01	-.24
			Spelling	+.14	+.05	+.14
			Language Mechanics	+.22	+.27	+.25
			Language Expression	+.11	+.05	+.23
Study 2: Lafayette Parish, Title I Math (JDRP #82-46)	Math, Grades 3-6 Chapter 1 students	Compared to students receiving Chapter I pullout. Randomly assigned.	CTBS Total Math: CAI > control	ES= .19		
Study 3: Merrimack Education Center (JDRP #82-34)	Reading, Grades 2-9 Chapter 1 students	Compared to students receiving Chapter I pullout. Randomly assigned.	Metropolitan Achievement Test, Reading: CAI > control	ES= .40		
Basic Literacy Through Microcomputers (microcomputers or typewriters used) (JDRP #84-14)	Reading, Grade 1 Chapter 1 students	Compared to students receiving no additional time. Randomly assigned.	California Achievement Test, Reading: Treated students > control	ES= .58		