The Effectiveness of Chapter I Pull-Out Programs on Reading Achievement.

The effects of Chapter 1 (Education Consolidation and Improvement Act) reading pull-out programs on the reading achievement of participating students was studied for 30 fifth- and sixth-grade students from a Chicago (Illinois) elementary school serving a low socioeconomic neighborhood with a minority enrollment of 99 percent. Fifteen students had received pull-out reading instruction, and 15 had not. Reading achievement changes were established by comparing scores on the Iowa Tests of Basic Skills (ITBS) from the end of third- and fourth-grades with scores from the end of fourth- and fifth-grades. Findings were tabulated in terms of means, medians, and standard deviations, and statistically significant differences between the mean scores were determined. Overall, the data indicate no significant difference in reading achievement between the pull-out and non-pull-out classes over one year. This study points to the need to continue such research using a larger sample and a longer span of time, particularly because other research has cast doubt on the efficacy of pull-out instruction. One table provides study findings. (Contains 25 references.) (SLD)
Since 1965, Title I/Chapter I of the Elementary and Secondary Education Act (ESEA) has been the bedrock on which federal aid to many urban school districts has been built. In many Chicago schools that serve a high proportion of disadvantaged children, such funding comprises a major budget item. During the 1990-1991 school year, federal Chapter I provided $5.4 billion to serve five million children. The 1991-92 school year saw this part of the national budget increased by sixteen percent to $6.2 billion. Furthermore, infusion of similar funding under state Chapter I will bring hundreds of millions of dollars to Illinois schools that enroll low-income pupils. Estimates for 1993-94 put such figures well over $300 million dollars.

Under provisions of the Chicago School Reform Act, the control and the designed use of state funds is being shifted from the central office to use by local school direction. Since 1988, the focus of federal Chapter I has also changed dramatically. Mandates now call for accountability for student performance, coordination with regular programs, and a need to creatively pursue performance results. It is therefore evident that Chapter I programming must be considered as an important implementation strategy for increasing the achievement of not only disadvantaged students, but also as a part of a total combined effort to raise the achievement level of the entire school's population. Discretionary control of such funds is meant to help schools meet performance objectives.

A number of schools in Chicago, such as the J. N. Thorp School discussed in this study, have traditionally provided Chapter I assistance for reading improvement via pull-out programs. Though this particular program design is not mandated as an intervention strategy, it is often utilized by schools since it comfortably "fits" the mandate that funding be used to supplement and not supplant the basic program. A program strategy in existence for a long time possibly risks becoming comfortable with standard operating procedures, bent on not rocking the boat, and dedicated to preserving the status quo.

After many years of observation at the national level, a large body of research now points to the need to carefully examine the effectiveness of program designs at the local level.
Though the use of pull-out programming has been in use for many years, existing methodology constantly needs reassessment. Under the direction of school reform, administrators, local councils, and teachers need to accurately measure pull-out methods to determine their effectiveness. We must use what we have learned from the past, from the constantly growing body of research on the disadvantaged, from literature on successful programs, and from current examination of the situation, if we are to make correct decisions which will effect achievement priorities in the future. Are students benefiting from this particular method of programming as it relates to reading achievement? Are such programs significantly increasing reading performance or should such funding be utilized in other directions which might be more suitable to a particular student body?

Given the proper information, teachers, administrators, and concerned parents can make informed decisions within their respective schools as to the merits of continuing past practices or the need to explore other avenues of instruction in the area of reading improvement. Since Chapter I funding represents, and may continue to represent, a large portion of the available supplemental funding in schools having a high incidence of disadvantaged students, the decisions to be reached will most likely have a far reaching effect on each particular school's student body, staff training, and the priorities given to school budgeting for the school's master plan to improve achievement and meet performance goals.

History of Chapter I (Title I)

The largest compensatory educational effort in the United States is the federally funded Title I of the Elementary and Secondary Education Act of 1965 (ESEA), now revised as Chapter I of the Educational Consolidation and Improvement Act of 1981 (ECIA). This funding was allocated to provide financial assistance to local educational agencies serving areas with high concentrations of children from low-income families to expand and improve their educational programs. Presently funded at $6.2 million, this delivery system utilizes twenty-two percent of the Department of Education budget (LeTendre, 1991). Funding is so widespread that nearly ninety percent of all school districts receive some funding, and in these districts nearly twenty percent of all elementary school students receive some services.

Of those students served by Chapter I programs, eighty-five percent are provided with instruction in reading or language arts for between two and one half (Allington, 1986), and three and one half hours per week, (National Institute of Education, 1977), the vast majority in "pull-out" compensatory instruction classes. After nearly three decades of Title I/Chapter I services, the amount of research literature both on its merits and weaknesses is
abundant. In this review I will attempt to address those research studies that appear to be most relevant to the question of the merits of Chapter I intervention as a vehicle for improving reading instruction.

National Studies

The Instructional Dimensions Study (IDS) was one of several studies conducted by the National Institute of Education (NIE) as part of the mandate of the 1974 Educational Amendments to study compensatory education supported by ESEA Title I. The study was based on data collected in the 1976 and 1977 school years (Cooley, 1980). This major study sought to assess the impact of a variety of classroom structures on reading and mathematics achievement. One major purpose of the study was to examine the differential effects of pull-out and in-class designs on the achievement of students who received compensatory aid. The data for the study were collected from 400 purposely selected first and third grade classrooms. This approach to sampling was designed to select classrooms that evidenced specific instructional dimensions, without specifically seeking to produce nationally representative results. Classrooms selected varied on student background factors such as ethnicity, family income, etc. Schools used in the sample were also selected to represent variations such as urban status, size, and percent of Title I enrollment. The final sample was made up of 400 classrooms in 100 schools from fourteen districts (Riddle, 1984).

The IDS had several strengths that made it relevant to an assessment of Chapter I's effectiveness. First, it attempted to primarily assess the effectiveness of compensatory education practices and did not focus on broader populations than the educationally disadvantaged. Secondly, the study went beyond mean effects to examine the differential effects of pull-out designs, most common to Chapter I. Thirdly, the study used the same measure of achievement in all classrooms, (CTBS, level B, form S and form S, level I, depending on grade tested) (Kennedy, 1986).

The IDS yielded three important findings (NIE, 1977; Cooley and Leinhardt, 1980). For the 400 first and third graders studied it found that individualized instruction did not make a difference in achievement.

Students gained the same amount whether taught individually or in groups. Secondly, it found that pull-out programs were a more effective instructional arrangement for some groups but not for others. First graders profited more from in-class programs. Third graders derived more benefit from pull-out programs in mathematics but showed no difference in reading. No attempt was made to project findings to other grades. A third important finding of this study was that the amount of time devoted to instruction, the size of the group, and the match between the curriculum and the content of the achievement test used as a measure, were more significantly related to student achievement (Linn, 1982).
The Sustaining Effects Study

The Sustaining Effects Study was another large scale attempt to analyze the effectiveness of Chapter I. This federally mandated investigation of compensatory education was conducted by Systems Development Corporation (SDC) and attempted to examine the "nature, quantity, and environment of compensatory education and its sustained effects" (Wang, and others, 1981). Funded at more than $20 million, the SES conducted a series of substudies between 1975 and 1980. The major thrust of this study was to examine the relationship between classroom practices and student achievement. Unlike the IDS study, data was gathered on instructional techniques from teacher questionnaires rather than from direct observations. The major outcome of interest was student achievement scores.

(Carter, 1980). The Sustaining Effects study attempted to examine the achievement progress of Title I and needy students in non-compensatory education schools relative to a nationwide sample. Sustaining Effects Study researchers concluded that the amount of compensatory instruction would not by itself close the achievement gap found between compensatory students and nondisadvantaged peers. Because researchers collected achievement data on all students in their sampled schools, they were able to compare achievement gains of Chapter I children to those of students who did not receive services. Central findings of the study conclude that percentile ranks of needy students not participating in compensatory education often declined. Comparison groups of fourth and sixth graders defined as needy but not receiving services rose in percentile rank over the year studied. Needy students not in Chapter I were at higher percentile ranks to start but did not improve as much as Title I students (Carter, 1983). Researchers also attempted to compare Title I students and needy non-Title I students to a representative sample drawn from 243 schools across the nation. To summarize the major findings: Students receiving Chapter I services experience larger increases in their standardized achievement test scores than "comparable" needy students who do not. However, their gains do not move them substantially toward the achievement levels of more advantaged students. Students in early elementary grade programs gain more than students participating in later-grades (Kennedy, 1986).

Critics of the above study argue the assumption that non-Title I students identified as "needy" were indeed "comparable" to Title I students since they were selected from school having relatively fewer poor students. More recent studies (Meyers, 1986) have tried to statistically correct for such differences in groups. Reanalysis of the SES Study data using analytic techniques indicated that the more similar the comparison group to Title I students, the greater the achievement benefits associated with Title I participation. However, patterns of effects found using the statistically optimum comparison group were not substantially different.
Title I/Chapter I Evaluation and Reporting System (TIERS)

In 1974, as part of the Educational Amendments Act of 1974, TIERS was developed by the U.S. Office of Education to evaluate Title I programs. In 1981, ECIA Chapter I repealed this authorization but mandatory data collection and reporting of evaluation data still remain for the years 1979 to 1981. Since 1981, collections and reporting of data have been voluntary. Analyses of TIERS data for the 1983-1984 school year have been evaluated (Carpenter & Hooper, 1985). Unlike other studies, TIERS does not depend on sampling techniques to estimate achievement gains attributable to Title I. Instead data, aggregated at various levels, are intended to assess gains in basic achievement and how they relate to Title I projects. Selection of achievement test is left to the discretion of states and local districts. Changes of achievement are reported in National Curve Equivalents (Linn, 1982, p. 9). National Curve Equivalents were then converted to percentile ranks. To summarize the findings: The achievement level for grades two through twelve indicated an upward gain for almost every grade. The size of the increase is often only a few percentile ranks, and Chapter I students' achievement after one year was still far below the median, or 50th percentile rank. In reading the gain averaged from 4 to 5 points for elementary aged students. Critics of the TIERS study point to many possible errors in statistical procedures. For example, students with lower test scores will often show larger growth than other students, a phenomenon known as "regression to the mean" (Meyers, 1986). Furthermore, gains of Chapter I students were compared not to comparable students, but instead were assessed relative to the entire score distribution. For results of this study to be meaningful, one would have to assume that students without Chapter I assistance would remain at their same relative percentile rank, a measure which is not particularly stable over time. Furthermore, one would have to assume that scores of non-participants are in fact comparable to Chapter I students having the same percentile rank (Linn, 1982).

Findings from both the TIERS and Sustaining Effects Study show that mathematics achievement yielded larger gains than reading achievement with the SES showing reading gains only in grades one through three. TIERS shows a pattern of decreasing gain as students move to higher grades (Advanced Technology, 1983).

Studies Relative to Service Delivery

Since volumes have been written by researchers dealing with the merits or faults of Chapter I, it is difficult to summarize the findings unless one concentrates on a specific area of concern. For example, Mullins, when completing a mega-analysis of Chapter I studies, reported no less than 5,000 references to compensatory education files in ERIC from the years 1970 through 1979 (Mullin, 1983). In his mega-analysis of 47 studies he found that the majority support the view that compensatory programs have a small
but positive effect on achievement growth of the disadvantaged. Their gains appear to be slightly larger than those predicated if they were not exposed to the programs. In his examination of studies concerned with service delivery, Mullins found that not only was there no significant agreement or stability to findings on the relative effectiveness of compensatory education in different grades, there was also no consensus as to which program type was most beneficial. William Cooley and Gaea Leinhardt (Coley, 1978) as part of a NIE evaluation of compensatory education tried to identify particularly effective practices and concluded that no one method—individualized or group—is superior.

The Design and Implementation of Chapter I Instructional Services

In 1986, as part of another national assessment of Chapter I, a study of 24 schools utilizing Chapter I programs was completed by the Far West Laboratory for Educational Research and Development (Rowan, Brian, 1986). Previous to this study a large amount of research had both criticized the wisdom of instruction via pull-out programs (Glass and Smith, 1977) (Kimbrough Hill, 1981). Some said such methods caused confusion with regular content classes (Allington, 1986). Other studies suggest that pull-out delivery, per se, has little effect on quality of instruction variables (Archambaut, 1986).

One of the questions that the Far West Study attempted to explore dealt with the need to analyze the design practices offered at various Chapter I sites. Most previous research had indicated that most school Title I projects used pull-out models (Glass and Smith, 1977), but others had used other designs such as in-class help and replacement methods (Advanced Technology, 1983). Did pull-out models disrupt ongoing lessons in regular classes as was suggested by some researchers (Kimbrough and Hill, 1981)? Did the implementation of pull-out designs result in a lack of coordination between compensatory and regular program instruction as suggested others? (Johnson, Allington, and Afflerbach, 1985). Since Chapter I legislation gave schools much latitude in design of service delivery, there was a need to investigate the implementation of various models (Cooper, 1986).

To summarize the results of the Far West Study is a difficult task. No one method of service delivery was shown to be more effective than another. The great variety in Chapter I "treatments" across schools makes generalizations about specific programs difficult. One uniform characteristic of Chapter I was found. All of the Chapter I projects offered instruction in small groups. Possibly past large scale evaluations of Chapter I have been influenced by this trend. Participation in smaller-sized classes has given Chapter I students an advantage over non-participants, however, the small effects found for Chapter I in "macro" evaluations make sense in light of this finding. Since Chapter I instruction usually accounts for less than 100 hours of learning time over the course
of an academic year, the effects of reduced class size on student achievement should not be large (Rowen, 1986).

The Far West Study also pointed out that researchers who argue that Chapter I programs steal valuable learning time from regular class instruction may be inaccurate in their estimations (Lignon and Doss, 1982). Instead of adding to the amount of time students spend in regular reading classes, most schools simply redistributed a fixed amount of instructional time across programs. This process, however, does not give Chapter I students additional time to learn. Studies by Kimbrough and Hill (1981) found that compensatory programs tended to replace core classroom instruction, especially in reading during the regular school day, even though many schools report no such loss due to pull-out arrangements (NIE, 1977).

Researchers such as Shulman and others believe that schools employing pull-out projects do so from reasons other than pedagogical concerns. Pull-out structures, he believes, produce a more easily followed "audit trail" enabling authorities to verify compliance with the "supplement but not supplant" regulations of Chapter I (Allington, 1986) (SRI, 1991).

Since the findings in a great number of the studies done on the effectiveness of various Chapter I designs seem conflicting except for the general agreement that positive results are being recorded; one can only point to the need for more specific evaluations. Further research needs to be done during the next decade from a more site specific point of view. Possibly the spirit fostered by the reform movement and technological advances made in the field of research will make such evaluations possible. This evaluation will be one attempt to make such an evaluation in a Chapter I urban school setting.

Therefore, the purpose of the study is to determine the effect of Chapter I reading pull-out programs on the reading achievement of participating students.

**Procedures**

**Population/Sample:**

The population of this study will include sixty fifth and sixth grade students from the J. N. Thorp Elementary School. This school, which has a grade range from pre-school to grade eight, is a Chicago public school located in the South Chicago area. The students reside locally in this predominantly low socioeconomic neighborhood. The total population of the school is approximately 780 students of which 99% are minority. Approximately 30% of the student body is Hispanic and 69% is African American. The mobility rate is 55.4% and the rate of attendance was reported as 91.8% on the 1991 State School Report Card. The average class size of the
regular classes used in this study for the intermediate grades was 21.7 students. The average number of minutes devoted to language arts instruction in regular classes at the grades in question is 103 minutes per day. The majority of students in the population (98%) studied come from homes receiving some form of economic assistance. Many are from single parent families. Over 50% of the teaching staff has been at the school over ten years, and the majority have advanced degrees in education.

For purposes of this study sixty fifth and sixth grade students presently enrolled at the school were selected. School records established that while in fourth and fifth grade respectively, thirty of these students received reading instruction in Chapter I pull-out reading laboratories. Fifteen students were selected at random from each of these sub-populations so as to create a control and experimental group of Chapter I and non-Chapter I reading pull-out recipients.

Each Spring, the Iowa Tests of Basic Skills (ITBS) are administered to each student in the Chicago Public elementary schools. Two sample groups of intermediate grade students were identified from the J. N. Thorp School records, namely those students who had received reading instruction in regular classes and those that had received reading instruction in a Chapter I reading laboratory. The grade level reading score on the ITBS obtained in the Spring of 1990 for each subject in the sample was compared to the reading level recorded for the Spring 1991 testing; thus allowing for a measure over a twelve month period. Including the scores of students from more than one grade made the inclusion of students from more than one classroom division, and those served by more than one Chapter I reading teacher, possible. A pre-posttest control group design was employed.

The findings were tabulated in terms of means, medians, and standard deviations. Calculations were made utilizing grade equivalent scores. Medians were reported as grade equivalent and as NCE scores for ease in comparison to other research studies. The t test for independent samples were utilized at the .05 level of confidence to measure any statistically significant difference between the mean scores.
Findings

The samples for the study included present fifth and sixth grade students at the J. N. Thorp Elementary School who were at the time of the pretest completing third and fourth grade, and were completing grades four and five at the time of the posttest. Each spring these Chicago public school students take the ITBS (Iowa Tests of Basic Skills). From these students two groups of fifteen students were randomly selected. Subjects in one group were taught reading in the traditional classroom setting while the other group received reading instruction in a Chapter I reading pull-out classroom. Results from the 1990 ITBS reading subtests were used as a pretest and results from the 1991 ITBS reading subtests were used as a posttest. A t test (p < .05) for independent samples was done for these sets of scores to determine if there was a statistically significant change in reading achievement due to Chapter I pull-out exposure. Table I summarizes the statistical analyses.

Table I

<table>
<thead>
<tr>
<th>Test</th>
<th>Experimental</th>
<th>Control</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.70</td>
<td>3.91</td>
<td>.144</td>
</tr>
<tr>
<td>Mdn</td>
<td>3.90</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>Mdn (NCE score)</td>
<td>42.00</td>
<td>45.00</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.665</td>
<td>.566</td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>4.67</td>
<td>4.74</td>
<td>.037</td>
</tr>
<tr>
<td>Mdn</td>
<td>4.40</td>
<td>4.90</td>
<td></td>
</tr>
<tr>
<td>Mdn (NCE score)</td>
<td>42.00</td>
<td>41.00</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.769</td>
<td>.950</td>
<td></td>
</tr>
</tbody>
</table>

* Significance at the .05 level
Examination of the 1990 mean pretest scores indicate that the Experimental and Control groups were not significantly different at the start: the E-group having as its grade equivalent mean a score of 3.70, and the C-group having a grade-equivalent mean score of 3.91. This finding allows for the assumption that the two groups were equivalent in reading achievement in the Spring of 1990.

Examination of the 1991 posttest scores indicates that after an additional year of school, and that year representing a full year of pull-out instruction for the E-group, the mean scores are 4.67 for the E-group and 4.74 for the C-group. Thus, no significant increase or decrease in reading achievement has occurred for either group. The t score for the 1990 results (.144) and the t score for 1991 results (.037) show no statistically significant change in reading achievement for the two groups.

It is interesting to note that when compared to the city wide median NCE (36) reported for all elementary students tested in 1991 by the Chicago Board of Education, students from J. N. Thorp showed slightly higher median NCE scores in 1991 (E-group = 42, C-group = 41). The median NCE nationally is 50, and the median NCE for all Chicago Chapter I students for reading in 1991 was 26 (Chicago, 1992).

Overall, the data from this study leads to the acceptance of the null hypothesis: no significant difference in reading achievement will be observed between pull-out and non-pull-out classes after one year of instruction. The observation tends to agree with Chicago Board of Education statistics which report that the median NCE gain in reading for all of Chicago's Chapter I students in the 1990-1991 school year was zero.

This study points to the need to continue such research utilizing a much larger sample and continuing the study to encompass a longer time span as students participate in compensatory programs for a longer length of time. Past research indicates that disadvantaged learners progress at a much slower rate. Further refinement of the study by stratified sampling may indicate different results for different age groups, students of different sex, or for students with different ethnic backgrounds.

The research already reviewed indicates, however, that pull-out instruction may not be providing the supplementary "additional" instruction that disadvantaged learners need to increase their level of achievement. More research is necessary to refine and improve the delivery of effective reading instruction.
References


