DOCUMENT RESUME

ED 390 929 TM 024 441

AUTHOR D'Agostino, Jerome; Hiestand, Nancy
TITLE Advanced-Skill Instruction in Chapter 1 Summer Programs and Student Achievement.
PUB DATE Apr 95
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Tests; Compensatory Education; Disadvantaged Youth; *Educationally Disadvantaged; *Elementary School Students; Grade 4; Intermediate Grades; Mathematics; Pretests Posttests; Reading Comprehension; *Skill Development; *Summer Programs; *Thinking Skills
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; Iowa Tests of Basic Skills

ABSTRACT

Summer school is thought to be an effective alternative delivery mode for Chapter 1 education. The effectiveness of summer Chapter 1 programs was studied during Fall 1993 in 68 Chicago (Illinois) public schools. Students who had been in fourth grade in 1991-92 were divided into a group that had received Chapter 1 help during the school year and had attended 1992 summer school programs (n=1,006) and a group that had only school-year instruction (n=490). Reading comprehension and mathematics scores on the Iowa Tests of Basic Skills were used as achievement pretest and posttests. Advanced-skill instruction during the summer, received by 407 of the students, was positively and significantly related to students' adjusted yearly math and reading gains. Simply attending summer school did not help the most educationally disadvantaged students catch their Chapter 1 peers. (Contains one figure, one table, and nine references.) (SLD)

Advanced-Skill Instruction in Chapter 1 Summer Programs and Student Achievement

Jerome D'Agostino

University of Chicago

Nancy Hiestand

Chicago Public Schools

Paper Presented at the Meeting of the American Educational Research Association, San Francisco, April, 1995

Advanced-Skill Instruction in Chapter 1 Summer Programs and Student Achievement

Attempting to close the achievement gap between educationally disadvantaged students and their more advantaged peers has been a fundamental purpose of ESEA Chapter 1 since its inception in 1965. There are a number of different Chapter 1 delivery models that schools can implement to serve their low-achieving students. Schools often use more than one model to serve students with varying needs at different grades, especially high-poverty schools (Millsap, Moss, & Gamse, 1993). Providing Chapter 1 over the summer months is one option available to schools. Approximately 15 percent of Chapter 1 schools nationwide allot funds for summer programs, but 28 percent of high-poverty schools do so (Millsap et al., 1993).

Summer school has the potential to be an effective alternative to more popular delivery models, such as pullout labs, because they do not supplant the regular-classroom instruction and cause students to miss important activities that occur during regular class time. Instruction during the summer is truly supplemental. Further, studies that have investigated student academic growth spanning more than one school year have found that most students' achievement growth rates slow down over the summer months (David & Pelavin, 1977; Gabriel, Anderson, Benson, et al., 1985), but it has been found that the rate of summer growth of disadvantaged students was substantially less than the rate of summer-growth of non-disadvantaged students (Heyns, 1986).

It seems reasonable that a well-developed and academically-oriented summer school program could serve to reverse this trend and allow disadvantaged students the opportunity to maintain or increase their academic status. Most summer school programs, however, are not academically rigorous (Heyns, 1986). Indeed, the value of Chapter 1 summer instruction has not gone unquestioned.

Results from the extensive Sustaining Effects Study, which was carried out during the late 1970's, indicated that students provided compensatory education services during the summer did not gain significantly more in either math or reading from spring to fall than students who did not attend summer school (Carter, 1984). These results were not surprising, because it was also found that little reading or math instruction occurred during summer in the sampled schools (Carter, 1984).

An attempt was made in the Sustaining Effects Study to document significant relationships between instructional components of summer school and student academic growth (see Klibanoff & Haggart, 1981, for details). Although no relationships were discovered, instructional activities often related to achievement gains were not assessed. Data were collected that described certain instructional arrangements, such as the degree of emphasis on individualized versus small group instruction, rather than more specific classroom activities. Furthermore, this information was collected by asking the teachers to provide self-reports.

Evaluations are needed to discover if students' academic growth over the summer depends on the quality of instruction they receive. This study was intended to partially address this issue. The relations between Chicago Public School Chapter 1 summer school instruction and fourth-graders' annual reading and math achievement gains were examined. This was not a typical process-product study, however, in which measures of fragmented classroom components are correlated with gains. The goal was to explore if students who were provided the opportunity to cultivate higher-order skills over the summer were more likely to post larger achievement gains than their peers who were not given such an opportunity.

Present Chapter 1 legislation encourages schools to emphasize advanced-skill or higher-order skills instruction for eligible students. Many studies have found that educationally disadvantaged students receive less training in these skills than their more privileged peers (Millsap et al., 1993). Traditional Chapter 1 instruction has focused mainly on basic-skill remediation, which is based on a widely accepted assumption that certain skills need to be mastered in a sequential and hierarchical manner. Although this instructional approach may work to improve students' basic skills, it has not had positive consequences on students' advanced-skill learning (Means & Knapp, 1991a). Studies have indicated that challenging disadvantaged students in the classroom is a more productive approach to improve their abilities to think constructively and develop problem-solving capacities (Means & Knapp, 1991b). In fact, this research also revealed that disadvantaged students learned basic skills as well or better when they were provided instruction that emphasized advanced-skill learning compared to conventional instruction.

Newmann and Wehlage (1993) have argued that advanced skills are best taught in an environment that fosters "authentic instruction," which provides students the opportunity to: use higher order thinking strategies; explore topics in depth rather than superficially; experience lesson related material that is connected to their lives; engage in substantive conversation with their peers and teachers; have social support in the classroom that promotes high levels of achievement. Instructional programs developed to teach students advanced skills usually incorporate some or all of these principles (see Means & Knapp, 1991a). Newmann and Wehlage's model served as a theoretical framework in which the effectiveness of advanced-skill instruction during the summer could be evaluated.

This study was undertaken in the Fall of 1993 and two general questions about the effectiveness of Chapter 1 summer school were investigated. These questions were:

1). Did summer school students exposed to instruction that was focused on advanced skills outperform summer school students without such exposure on the Reading Comprehension and Math Total scores of the Iowa Test of Basic Skills (ITBS)?

2). How did summer-school students perform compared to regular-year-only Chapter 1 students, and did differences in performance depend on the level of advanced-skill instruction offered during the summer?

Method

Sample and Data Sources

Chapter 1 students who either attended or did not attend summer school were located in a sample of 68 Chicago Public Schools. All students were enrolled in the fourth grade during the 1991-92 school year. Two groups of students were identified. The first group, which was considered the treatment group, was those students who received Chapter 1 during the school year and who attended 1992 summer school programs. The second group (the control group) was composed of students who received regular-year Chapter 1 (during the 1991-92 school year) but did not attend summer school.

Classroom observations had been made during the summer of 1992 in the 68 schools. The schools were sampled at random from a population of about 250 Chicago Public Schools that received Chapter 1 dollars. District enrollment form data identified the students who were registered in each observed classroom. Summer school observations were available from 90 fourth-grade rooms in the sampled schools.

Observation Form

The observation system required observers to rate on a four-point scale the quality of 15 various activities that may have occurred in each summer-school classroom. If the activity did not occur during the observation period, a zero was coded. Observers also were asked to rate on a four-point scale the frequency of seven dimensions related to classroom climate.

In order to answer the second research question specified above, it was necessary to create a quantitative measure of students' experiences with advanced-skill instruction. Newmann and Wehlage's concept of "authentic instruction" was used to guide the creation of the variable. Although the summer observation device was not designed specifically to measure the level of "authentic instruction" occurring in classrooms, many of the items represented certain principles of the construct. Seven of the 15 activity items and four of the seven classroom climate items were used to represent the factor. The items were:

Activities

A. Teacher introduces a concept or defines a term

B. Teacher promotes problem solving

C. Teacher promotes critical thinking

D. Teacher promotes creative thinking

E. Teacher extends questioning/probing to broaden knowledge

F. Teacher promotes cooperation or teamwork among students

G. Teacher allows for active student participation in the learning activity

Classroom Climate

H. Students demonstrated on-task behavior

I. Students felt free to ask questions or obtain help

J. Students appeared to understand/comprehend materials

K. Teacher expressed high expectations for student learning

The Rasch measurement model was used to create an Authentic Instruction variable. The procedure and results are explained in the Results section below.

Observation Procedures

Seven individuals observed for approximately 45 minutes in each fourth-grade summer classroom at the sampled schools. Observations were made of either reading or math lessons, and they were not made in classrooms where students were taking tests or watching films. A reading lesson had to consist of students either reading or listening to some form of a written passage. A lesson focused solely on language arts that did not include a segment where material was read by students or another person was not deemed admissible as "reading." Observers completed the observation form immediately following the observation period.

Test Scores

1992 and 1993 ITBS Reading Comprehension and Math Total scores in Normal Curve Equivalents (NCEs) were collected for each fourth-grade Chapter 1 student enrolled in the sampled schools. The 1992 scores served as the pretest and the '93 scores served as the posttest.

Results

Development of the Authentic Instruction Variable

The Rasch measurement model was imposed on the data (see Note 1 for an explanation of Rasch measurement) in order to: (1) validate the meaning of the Authentic Instruction variable; and (2) to compute a measure on the variable for each classroom. Figure 1 is a map of the variable which contains both classrooms and the authentic instruction principles. The measure values in logits are located in the far left column. They are ordered from largest at the top of the figure to smallest at the bottom. The larger the measure on the variable, the greater the emphasis on authentic instruction in the classroom. As can be seen, the column containing asterisks represents the distribution of classrooms on the variable. A classroom's measure in logits is its natural log odds of receiving a mid-scale rating (i.e., about 2.5) on items that define the zero point on the scale. The original scale values are located in the column on the far right of the map.

The Rasch analysis revealed that no item outfit value was less than two, which indicated that all 11 items fit together sufficiently well to create the variable. A measure on the variable was computed for each observed classroom. Some of the classrooms were detected as misfitting, which indicated that the observers did not score these classrooms in the same pattern as the others. For these classrooms a valid measure on the variable could not be estimated precisely. The classrooms that misfitted the model were not included in further analyses. There were 79 classroom measures retained and eleven were dropped.

Summer School Instruction and Achievement

The first series of analyses were performed to investigate the relationship between the students' achievement gains in math and reading and their level of exposure to advanced-skill instruction during summer school. Hierarchical Linear Modeling (HLM) was used to examine these relationships (see Note 2 for a description of HLM).

In order to control for a regression-to-the-mean effect, adjusted math and reading gain scores were computed for each student. Posttests were regressed on pretests and the residuals were considered the adjusted gains. Thus, the residuals were gains corrected for pretest status.

Two means-as-outcomes HLM analyses were performed: one using the adjusted reading gains and one using the adjusted math gains as outcome variables. The Authentic Instruction measure was used as the predictor variable.

Authentic Instruction was significantly and positively related to students' adjusted reading gains, $t = 2.03$, $p < .05$. Also, the factor was significantly and positively related to adjusted math gains, $t = 2.04$, $p < .05$. Summer classrooms that offered a greater focus on learning advanced skills contained students who made greater yearly reading and math achievement gains.

Achievement of Chapter 1 Summer School Participants Compared to Non-Participants

In order to compare the achievement gains of students who attended summer school to those who did not, each student was categorized into one of four groups. Three of the groups were created based on the level of Authentic Instruction they received during summer; low, moderate, or high. The low group of classrooms were those that were below the 34th percentile on the Instruction measure, the middle group were classrooms between the 34th and 64th percentile, and the high group were the top third of the classrooms. Students who attended summer school were categorized into one of the three groups based on the category in which their summer school classroom was placed. The fourth group of students served as a control group, or those that received regular-year Chapter 1 services but did not attend summer school.

Returning to Figure 1, it is possible to see the types of principles that best characterized the summer classrooms based on the three levels of Authentic Instruction The low group were those rooms below -.50 on the measure and the cutoff separating the moderate from the high groups was 0 (some of the classrooms that appear to have a 0 measure actually had measures either slightly above or below 0). A classroom with the same measure as a principle had a 50 percent chance of receiving an average rating from the observer on that given principle. The odds of receiving an average rating on principles decrease as the principle measures increase above the measure of the classroom. Likewise, as principle measures decrease below the measure of the classroom, the odds increase that the classroom received an average rating on the principles. Thus the activities or features with equivalent measures or lower than the classroom measure were those activities that best characterized instruction in that room.

The low group of classrooms were settings where students basically remained on-task, moderately understood the lesson, felt free to ask questions, and were active participants in the lesson. These classrooms were not settings where students were encouraged to engage in higher-order thinking. The moderate-level rooms were places where all of the features of the low group were intact and where teachers had high expectations for learning, introduced and developed new concepts, and extended questioning of students to broaden their knowledge. The high-level classrooms were places where all of the features of the first two groups were evident and teachers promoted student cooperation, critical and creative thinking, and developed student problem-solving skills. Thus, it was this group of classrooms that primarily focused on what is commonly considered advanced-skill instruction.

Two analyses of covariance were performed to compare reading and math test scores across the four groups. Controlling for pretest scores was especially necessary in this portion of the study because it was assumed that children chosen for summer school had lower pretest scores than their peers who were not chosen. In order to adjust for differences in the four groups on pretest scores, the 1992 ITBS scores were used as covariates. The 1993 ITBS scores were the outcome variables. Table 1 shows the test score means and standard deviations of the four groups.

Adjusted 1993 reading scores varied significantly across the four groups, $F = 2.80$, $p < .05$, as did adjusted 1993 math scores, $F = 5.18$, $p < .05$. In order to discover which groups significantly differed, follow-up contrast effects were performed.

For reading, the control group students had significantly greater scores than students who went to summer school (i.e., students in the three treatment groups combined), $F = 6.5$, $p = .01$. Students in the high group, however, did not score significantly lower (after controlling for 1992 scores) than students in the control group, nor did students in the low group.

For math the trends were more pronounced. After adjusting for 1992 scores, the control group out-performed the three treatment groups taken together, $F = 8.00$, $p < .01$. Only the high group, however, had scores that were equivalent to the control group.

Discussion

Advanced-skill instruction during the summer months was positively and significantly related to fourth-grade Chapter 1 students' adjusted yearly math and reading gains. But summer school students' gains were significantly smaller than those made by fourth-grade Chapter 1 students who did not attend summer school. Simply attending summer school did not help the most educationally disadvantaged students catch their Chapter 1 peers who did not attend school in the summer.

Students who were exposed to summer-school instruction that focused on advanced-skill learning, however, made slightly larger math gains than control students and approached the control group's level of math achievement (after controlling for pretest status). Students exposed to moderate and low levels of advanced-skill instruction during the summer fell further behind their peers.

Grouping students into three levels of advanced-skill instruction for comparison with control students led to inconclusive reading results. Control students out-performed all summer students, but did not out-perform either the low or high advanced-skill groups. Apparently, summer school instruction focusing on advanced-skill learning has more distinguishable effects on students' math achievement.

Summer school instruction needs to be improved if students most in need will have the chance to catch the actual achievement levels of their peers. Clearly, the findings indicated that merely encouraging students to remain on-task, getting them to understand and be involved in the lesson, and creating an atmosphere of acceptance during the summer are not enough to significantly improve their standardized math and reading achievement scores. These features are important, but they are not sufficient. Summer school classrooms also need to provide lessons that emphasize higher-order-thinking strategies if they are to be effective in improving the achievement of the most educationally deprived Chapter 1 students.

Unfortunately, however, observers reported that few classrooms highly emphasized problem-solving, creative and critical thinking, and teamwork. Even the highest measured classrooms in the sample were not settings where these activities were frequently observed. This can be seen in Figure 1 where it is evident that few classrooms had measures that were equivalent to or greater than the measures of these activities. Perhaps summer school can be more effective if these activities become the predominant focus of summer instruction.

The main methodological weakness of this study pertained to the observation procedures. It was not possible to study the inter-rater reliability because raters never observed the same classroom simultaneously. Further, raters observed only one lesson per classroom in either reading or math so it is questionable if they were adequately exposed to each classroom to render sufficiently accurate judgments. This was probably most likely in classrooms where lessons were brief. It would have been preferable if raters observed over a series of days in each classroom and if they observed both reading and math lessons in each classroom to control for differential teaching styles across the two subjects.

These problems probably inflated the measurement error, which rarely increases the probability that a Type I error will be made. It usually delimits the amount of power available for analyses. Consequently, the impact of advanced-skill instruction during summer school on participating students' achievement gains may have been more pronounced than the results of this study indicated. Nonetheless, further studies need to be based on more thorough classroom observation procedures that yield more accurate measures of instructional practices in order to better our understanding of the value of summer schooling.

References

Bryk, A.S., & Raudenbush, S. (1993). Hierarchical linear models: Applications and data analysis methods. Newbury Park, CA: Sage.

Carter, L. (1984). The Sustaining Effects Study of Compensatory and Elementary Education, Educational Researcher, 13, 4-13.

David, J. L., & Pelavin, S. H. (1977). Research on the effectiveness of compensatory education programs: A reanalysis of data. Menlo Park, CA: Stanford Research Institutes.

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

Klibanoff, L.S. & Haggart, S.A. (1981). Summer growth and the effectiveness of summer school. Technical Report #8 from the Study of the Sustaining Effects of Compensatory Education on Basic Skills. Santa Monica, CA: System Development Corp. (ERIC Document Reproduction Service No. ED 213 789).

Means, B. & Knapp, M. (1991a). Cognitive approaches to teaching advanced skills to educationally disadvantaged students, Phi Delta Kappan, 73, 282-289.

Means, B. & Knapp, M. (1991b). Teaching advanced skills to educationally disadvantaged students. Washington, DC: U.S. Department of Education.

Newmann, F.M. & Wehlage, G.G. (1993). Five standards of authentic instruction, Educational Leadership, 50(7), 8-12.

Wright, B. (1977). Solving measurement problems with the Rasch model, Journal of Educational Measurement, 14, 97-116.

Notes

Note 1

Rasch measurement transforms raw scores into linear measures for people and items. It does this in a three-step procedure. In the first step an initial person measure is computed by taking the natural log of the person's raw score on all of the items divided by the maximum possible raw score minus the person's raw score. This log computation is known as a "logit." Thus, in this study the maximum raw score for each classroom as rated by observers on the 11 items was 51. There were five possible points for each of the seven "activities" item (the four-point rating scale and zero) and four possible points for each "climate" item. In the second step, an initial calibration is computed for each item by using the same procedure performed in Step 1 for the persons. Step 3 involves the computation of an adjusted logit for each person and item based on the variability of the initial person measures and item calibrations.

Once the final logits are computed for the people and items, fit statistics can be calculated for each person and item. These fit statistics indicate which items do not adequately represent the developed variable and for which individuals a valid measure on the variable is not possible.

An explanation of the logic undergirding the Rasch model is necessary to understand the technique and how it was applied in this study. The item calibrations in logits indicate the ordering of the items on the created variable. Across the classrooms, observers found it easier to give higher ratings on items with lower calibrations. Items with large calibrations were those that observers did not rate leniently across rooms. In other words, the larger the item calibration, the more difficult it was for the teacher to perform the activity (or establish the climate) in a highly proficient manner.

A larger classroom measure in logits indicated a greater degree of emphasis on the created Authentic Instruction variable. Further, the greater the classroom measure, the more likely the classroom was rated high on any given item. However, if a classroom was rated high by an observer on more difficult items, but low on easier items, the fit statistic for that classroom would have been large, indicating that the observer of the classroom did not rate it congruently with the other classrooms.

Likewise, if a number of higher-rated classrooms received lower ratings on easier items than lower-rated classrooms, the item would have been detected as a misfit, which would indicate that the item did not represent the intended construct (i.e, "authentic instruction"). See Wright (1977) for a more thorough explanation of Rasch analysis.

Note 2

HLM is a series of regression methods that are useful for the analysis of multi-level data. The data in this study were considered multi-level because students, which were the level one unit, were nested within classrooms, which was the level two unit. One type of HLM model is the means-as-outcomes regression, which involves predicting the mean values of the outcome variables at Level 1 (student test score gains) from the Level 2 variable (Authentic Instruction measure). A t-value is computed, and when significant, indicates that a level two predictor variable accounts for significant variability across Level 2 units (i.e., classrooms) in the Level 1 outcome. Bryk and Raudenbush (1993) provide a comprehensive description of HLM.

Figure 1: Map of Classrooms and Principles of Authentic Instruction Factor

```
------------------------------------------------------------------------------------------------------
|Measure|+Classrooms   |-Principles                                                          |S.1   |
------------------------------------------------------------------------------------------------------
+      2 +              +                                                                    +(4)   +
|        |              |                                                                    |      |
|        |              |                                                                    |      |
|        |              |                                                                    |      |
|        | .            | CREATIVE THINKING                                                  | ---  |
+      1 + .            + PROBLEM-SOLVING                                                    +      +
|        | *            | CRITICAL THINKING                TEAMWORK/COOPERATION              |      |
|        |              |                                                                    | 3    |
|        | **           |                                                                    |      |
|        | ****+        | HIGH EXPECTATIONS                                                  | ---  |
*      0 * *******.     * T. EXTENDED QUESTIONING          T. INTRODUCED CONCEPT             *      *
|        | ******       |                                                                    | 2    |
|        | ****.        |                                                                    | ---  |
|        | ******.      | ACTIVE STUDENT PARTICIPATION     STUDENTS FREE TO ASK QUESTIONS    | 1    |
|        | ****         |                                                                    |      |
+     -1 + *.           +                                                                    + ---  +
|        | .            | STUDENTS UNDERSTOOD MATERIAL                                       |      |
|        | .            |                                                                    |      |
|        | .            | STUDENTS ON-TASK                                                   |      |
|        |              |                                                                    |      |
+     -2 +              +                                                                    +(0)   +
------------------------------------------------------------------------------------------------------
|Measure| * = 2, . = 1  |-Principles                                                         |S.1   |
------------------------------------------------------------------------------------------------------
```

Table 1: ITBS NCE Means and Standard Deviations by Group

Group	Reading 1992	Reading 1993	Adjusted Reading 1993	Math Total 1992	Math Total 1993	Adjusted Math 1993
Control (N=490)	28.12 (13.80)	31.48 (14.64)	29.88	27.36 (17.77)	28.22 (15.60)	26.49
Low Authentic Instruction (N=284)	24.14 (11.70)	27.68 (12.34)	28.31	24.38 (15.54)	23.89 (14.94)	23.23
Moderate Authentic Instruction (N=315)	24.57 (11.51)	27.13 (12.37)	27.41	21.74 (14.19)	22.88 (13.70)	24.47
High Authentic Instruction (N=407)	22.47 (11.07)	27.32 (12.50)	28.71	22.70 (15.23)	25.15 (14.22)	25.98

Note: First entry in each cell is the mean. The second entry (in parentheses) is the standard deviation. The adjusted test score standard deviations are the same as the unadjusted 1993 standard deviations.