DOCUMENT RESUME

ED 333 022 TM 016 501

AUTHOR	Holdzkom, David; And Others
TITLE	Better Teaching for Better Learning: Student Achievement Results in a 4-Year Pilot Teacher Career Development Plan.
PUB DATE	Apr 90
NOTE	39p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, March 16-20, 1990).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Achievement Gains; Achievement Tests; *Career Development; Elementary Education; *Elementary School Students; Elementary School Teachers; Grade 3; Grade 6; Grade 8; Pilot Projects; Program Effectiveness; Program Evaluation; Public School Teachers; *State Programs; *Teacher Improvement; *Test Results
IDENTIFIERS	California Achievement Tests; *North Carolina Career Development Program

ABSTRACT

To assess the impact of the North Carolina Career Development Plan (CDP) on student achievement, results of the California Achievement Tests, administered annually in North Carolina public schools, were analyzed. The focus was on the performance of children in grades 3, 6, and 8 in the school years from 1985-86 through 1988-89 in 16 CDP units. Performance data were also analyzed for students in 15 matched units not participating in the CDP in order to isolate patterns in performance data that could not be attributed to chance or other reform efforts. When student achievement in CDP units was examined, the general tendency was toward improvement. Fewer CDP units in grades 3 and 8 scored below the national median in 1989 than in 1986. In the group of matched units, the same trend was apparent only for grade 3. Improvement was more likely for CDP units than for non-CDP units. In grade 3, 13 CDP and six non-CDP units improved; in grade 6, 12 CDP units improved or showed no change and eight non-CDP units improved; and in grade 8, 13 CDP and eight non-CDP units improved. Some of this improvement can be attributed to the CDP, although it is not clear how much gain in achievement can be allotted to CDP participation. Sixteen tables present achievement data. An 18-item list of references is included. (SLD)

BETTER TEACHING FOR BETTER LEARNING: Student Achievement Results in a 4-Year Pilot Teacher Career Development Plan

by

David Holdzkom, Chief Consultant

Barbara Kuligowski, Consultant

Dennis Stacey, Consultant

Division of LEA Personnel Services
North Carolina Department of Public Instruction
Raleigh, NC

Presented at the 1990 Annual Conference of the
American Eductional Research Association, Boston, MA

INTRODUCTION

For several years, the education press has been filled with discussion and debate about the desirability/advisability/possibility of evaluating teachers' performance on the basis of observed behavior. Originally, much of the discussion focused on criteria. While there seemed to be general agreement that good teaching was recognizable, there arose general disagreement about the component skills that, taken together, equated to "good teaching". An interesting subset of this debate concerned the fundamental classification of teaching as either a science (in which case, similar effects could be expected from similar stimuli) or an art (in which case, the idiosyncrasies of individual artists prevented valid generalization from case to case.) While this debate continued, policy makers, ignoring the niceties of scholarly discussion, changed the stakes considerably.

Beginning in the early 1980's, legislatures, boards of education, and other stake-holders began to commission teacher evaluation systems, some of which would result in personnel decisions: granting tenure, offering salary increases, and certification. At this point, the education press began to feature debate about purposes of evaluation, not just criteria.

While the two debates, which may be viewed as reciprocals, have not been resolved, it is important to recognize that the heart of both issues concerns the relationship of teaching to learning. What is it that teachers do that causes students to learn? Most participants in the discussion are willing to forego causative relationships, but surely there exist correlations. Indeed the research literature for the last 15 years has emphasized an influential body of studies (loosely referred to as the "process-product" research) that posits such correlative skills and outcomes. (Gage and Needels, 1989; Brophy,

1988; Gage, 1984; Rosenshine, 1982). Moreover, in one of the few empirical studies of teachers in effective schools, Teddlie, Kirby, and Stringfield demonstrated that teachers in "effective schools" demonstrated more effective behaviors than did teacher in the ineffective schools. (Teddlie, Kirby, and Stringfield, 1989).

In 1985, the North Carolina General Assembly authorized the implementation of a pilot Career Development Plan. The pilot involved sixteen school districts from across the state, chosen to represent a variety of sizes, geographic distribution, and relative wealth. The innovation was intended to test the effects of differential rewards and opportunities for teachers on attraction to and retention in the profession. While improved student achievement was not a specific project goal, the use of a teacher evaluation system based on the process-product research was expected to have a positive impact on students.

During the first year of the pilot, this new performance appraisal system was installed in sixteen school districts in North Carolina (Holdzkom, 1987). This new appraisal system as implemented in the pilot units differed in several critical respects from earlier efforts:

1. The criteria for the first five functions, related to in-class performance, were drawn from the research literature in which the link between student achievement and the criteria was established.

2. The evaluation processes were standardized across all user sites (State Board of Education, 1986).

3. The evaluation data were gathered by at least two independently functioning raters.

4. All raters had received training in the criteria for evaluation Effective Teaching Training and in the processes of evaluation Teacher Performance Appraisal System.

Several program evaluations have reviewed the Teacher Performance Appraisal System in the five years since its implementation. An independent review of the performance appraisal system concluded that it was " a system admirably suited to its purposes", but cautioned that the system probably could not be

used to make evaluative judgements beyond Career Status II (Brandt et. al, 1988)

A study conducted by staff of the Division of Personnel Relations, North Carolina Department of Public Instruction, concluded that users of the system agreed that the criteria were appropriate for evaluating teachers and that, generally, evaluators possessed the skills and knowledge needed to carry out the evaluation process (Stacey et. al, 1989). Finally, a second third party evaluation found that evaluatees generally agreed with ratings given them and that the system was superior to the evaluation system it replaced (Research & Service Institute, 1989).

While improved instruction and consequently improved learning were clearly desirable, the interventions directly affected teachers, not students. Whereas curriculum changes directly influence, or seek to influence, the learner, the Career Development Program aims at instruction--or teacher behaviors--as the locus of change. While the effect on student achievement is desired, it is not directly traceable to CDP features.

This indirect relationship between strengthening instruction and strengthening student performance is further complicated by the fact that the specific teacher behaviors that are encouraged are themselves correlative, not causal, with improved student achievement. In the research that underlies North Carolina's performance appraisal system, all of the teaching skills are correlated powerfully with student achievement gains. It is, therefore, reasonable to ask: What effect has the Career Development Program pilot had on student achievement in the 16 participating districts? This study seeks to answer that question.

Frankly, we undertook this analysis with some reluctance. Throughout the pilot period, superintendents, principals, and teachers reported almost universally that teaching had improved and that student performance had improved. In 1986, we reported that 57% of more than 4,000 teachers responding to a Department of Public Instruction survey agreed that "participating in the CDP has helped me perform my role more effectively". (N.C. Board of Education, 1986). Only 31% of respondents disagreed with the statement.

In addition, 43% of respondents agreed that the Effective Teaching Training program and the new evaluation system had improved the quality of instruction in their schools. An additional 28% weren't sure. Even at that early date, 41% of respondents agreed that the Career Development Plan was likely to influence education in a positive way. Of all respondents, only about one-third (36%) disagreed with that statement.

Moreover, in a survey conducted by the North Carolina Association of Educators at about the same time (NCAE, 1986), 37% of respondents felt they had become better teachers and that classroom instruction had somewhat improved. Another 13% felt that they had definitely become better teachers. A quarter felt there had been no change, and only 17% felt classroom performance had suffered. In Spring, 1988, in a second NCAE survey (Bunche, 1988), 58% of the respondents agreed that "lessons based on the 6-step lesson plan have improved student learning", while 64% agree that "observation and evaluation have helped me to improve specific aspects of my teaching".

Our reluctance to undertake this analysis does not stem from any fear that student achievement suffered because of the Career Development Program pilot. Rather, we hesitate because of the complex interaction between teacher inputs and student outcomes. Teaching and learning are far too complex to reduce to a single number. Much of what is taught and learned goes untested. That is in the nature of evaluations which are <u>per force</u> limited in what they can measure.

Moreover, students experience far more than is taught by teachers. Put another way, teachers exercise too little control over their students' total experience for student achievement tests to be valid in discovering the relationship between any single teacher's efforts and any single student's achievement. Even in elementary schools, teachers instruct children for less than six hours, or 25%, of the day. At the high school, the percentage of contact hours between a given teacher and his/her students is even lower. However, if we consider unit average test scores, and if we consider this analysis as programmatic evaluation, much of the difficulty disappears.

We will still not factor out all the interference between teaching and learning, but by aggregating the data at the unit level and then comparing units, we will gain a useful sense of change in performance. Moreover, if we find trends or patterns in the data, these will be highly suggestive. In short, we should consider our data as analogous to the Dow Jones average. By monitoring change--either up or down--in a group of stocks, economists can make estimates of the strength of the national economy. The fact that one or two issues within the Dow Jones group behave differently from the other issues does not negate the value of the group performance. Similarly, in our report, the anomalous behavior of one or two school units, in one or two grade levels, may be cause for concern by district executives. Such behavior, however, does not negate our ability to make general statements about performance trends of the total group of school districts.

With these limitations in mind, then, we proceed to the actual analysis of the data. We do not believe that these data will indicate causal relationships between Career Development effects and student achievement. We do, however, believe that we will see correlations between participation in Career Development and improved student achievement. We cannot attribute all positive change to Career Development, but, if we fail to discover similar patterns of achievement in a matched sample of students, then clearly some effect of Career Development is present and is influencing students' achievement.

METHODOLOGY

In order to assess the impact of Career Development on student achievement, we analyzed the results of the California Achievement Tests (CAT), which are administered annually to children in North Carolina's public schools. Because of the availability of data, we focused on the performance of children in Grades 3, 6, and 8 during school years 1985-86, 1986-87, 1987-88, and 1988-89. Because the CAT was re-normed after the 84-85 administration, comparisons prior to the pilot period were not attempted. We also collected and analyzed performance data for youngsters in 15 additional units. We hoped, by using these comparison analyses, to be able to isolate patterns in the performance data that could not be attributed to chance or to other reform efforts--most notably the Basic Education Program and the Standard Course of Study--that have been undertaken in North Carolina schools.

The match units were selected by Dr. Carol Furtwengler of the Research & Service Institute of Brentwood, Tennessee. Dr. Furtwengler was the third-party evaluator retained by the N. C. General Assembly to study CDP implementations. She selected the match units on the basis of geographic distribution, average daily membership in the match units, per-pupil expenditure (excluding school food service), and percentage of students planning to attend college. (RSI, 1989). These variables are used because they are predictive of student achievement. Thus, Dr. Furtwengler constructed a mirror-image of the CDP pilot units. Similarities and differences in performance profiles in the pilot and match units should, therefore, allow us to draw some tentative conclusions about the effects of Career Development. Since the match units did not voluntarily participate in this study, we will refer to them by letter (District A, etc.) in order to avoid any hint of criticism of the performance of their students or teachers. (Research and Testing Services, 1987, Division of Testing, 1988).

RESULTS

The CAT scores for third grade students in the pilot units are presented in Table 1. For convenience, we have selected the median national percentile average on the total battery for each unit as our reported figure. For each year, we have arranged the units in rank order, followed by the unit's average score. The number in parentheses following the 1987 and 1988 scores indicates the amount of change from the prior year. Following the 1989 average, we present two numbers in parentheses. The first shows the change between 1988 and 1989. The second indicates the amount of change during the entire pilot period, 1985-86 and 1988-89. The plus sign (+), of course, indicates gain, while the minus (-) indicates loss. The dark line within the columns separates units scoring above the 50th percentile, or national average, from those units scoring below the national average.

Grade 3 Results

Examination of Table 1 makes several things clear. First, the number of CDP units scoring below the 50th percentile declined steadily over the period from four to two to one. While one unit continued below the median in 1988-89, it was within one percentile point of attaining the median. Moreover, both the ceiling (highest score) and floor (lowest score) rose over the period, when the units are taken as a whole. For individual units, the averages increased in 13 units over the period, and declined in three units. In the units experiencing decline over the period (Haywood, Harnett, and Alexander), however, two of them had relatively modest losses of two percentile points.

The gains made by the four units performing below the 50th percentile in 1985-86 should not be overlooked. These units all moved up significantly over the period, with Perquimans registering a dramatic 18 percentile point increase.

TABLE 1

3rd Grade CAT Scores (Unit Average) for CDP Units

1986		1987		1988		1989	
Burlington	76	Burlington	73 (- 3)	Burlington	77 (+ 4)	Burlington	80 (+3; +4)
Haywood	71	Tarboro	70 (+ 9)	Haywood	71 (+ 3)	N. Hanover	70 (+2; +9)
R. Rapids	67	R. Rapids	68 (+ 1)	Alexander	71 (+12)	Haywood	69 (-2; -2)
Buncombe	67	Haywood	68 (- 3)	Perquimans	70 (+ 8)	Buncombe	69 (+1; +2)
Harnett	64	N. Hanover	67 (+ 6)	Burke	69 (+ 6)	R. Rapids	69 (+6; +2)
Burke	64	Buncombe	65 (- 2)	N. Hanover	68 (+ 1)	Burke	67 (-2; +3)
N. Hanover	61	Burke	63 (- 1)	Buncombe	68 (+ 3)	Perquimans	66 (-6; +18)
Tarboro	61	Perquimans	62 (+14)	Tarboro	67 (- 3)	Mecklenburg	66 (+8; +6)
Mecklenburg	60	Alexander	59 (N.C.)	R. Rapids	63 (- 5)	Tarboro	65 (-2; +4)
Orange	59	Montgomery	59 (+ 7)	Orange	62 (+ 5)	Montgomery	63 (+7; +11)
Alexander	59	Mecklenburg	57 (- 3)	Harnett	60 (+ 4)	Orange	62 (NC; +3)
Montgomery	52	Orange	57 (- 2)	Mecklenburg	58 (+ 1)	Greene	59 (+4; +12)
Perquimans	48	Harnett	56 (- 8)	Montgomery	56 (- 3)	Harnett	58 (-2; -6)
Greene	47	Salisbury	52 (+ 8)	Greene	55 (+ 9)	Alexander	57 (-14; -2)
Salisbury	44	Greene	46 (- 1)	Chowan	50 (+11)	Chowan	52 (+2; +9)
Chowan	43	Chowan	39 (- 4)	Salisbury	47 (- 5)	Salisbury	49 (+2; +5)

TABLE 2

Range and Decile Distributions of CDP Units Per Year

	Range	Deciles : Units
1986	43 - 76 (33 points)	30th: 0; 40th: 4; 50th: 3; 60th: 7; 70th: 2
1987	39 - 73 (34 points)	30th: 1; 40th: 1; 50th: 6; 60th: 6; 70th: 2
1988	47 - 77 (30 points)	30th: 0; 40th: 1; 50th: 4; 60th: 7; 70th: 4
1989	49 - 80 (31 points)	30th: 0; 40th: 1; 50th: 4; 60th: 9; 70th: 1; 80th: 1

Moreover, of nine units scoring at or above the 60th percentile in 1985-86, seven units had improved by 1988-89. Put another way, units in all parts of the range were improving performance. Table 2 shows the difference per year in the range of scores and the distribution of units to deciles.

In summary, then, 81% of the CDP units showed improvements over the period, in performance of 3rd grade students, while 19% of the units showed decline. Overall, both high- and low-achieveing units experienced improvement, with only one unit remaining below the 50th percentile at the end of the period.

Tables 3 and 4 display the same information relative to the match units. It is clear that achievement in the match units, as a group, is somewhat lower than the pilot units throughout the period. Four of the match units are below the 50th percentile in 1985-86 and 1986-87, three remain below the 50th percentile 1987-88, and five units are below in 1988-89. Three of these units have been below the 50th percentile throughout the period, although the floor rises from the 31st to the 42nd percentile. Moreover, the ceiling rises from 64 to 69, although Unit L rose from the 49th to 72nd percentile before settling down to the 53rd percentile in 1988-89.

Overall, six units witnessed improvement in 3rd grade, while nine experienced declines or remained unchanged. Two of the four units that began below the 50th percentile experienced gains and two remained unchanged. As Table 4 shows, the range decreased over the period, indicating that growth at the top is occurring more slowly than growth at the bottom of the group.

In summary, then, the match units as a group show some of the tendency towards improvement seen in the CDP units. The growth, however, appears to be slower. As analysis of the individual units makes clear, the growth is, however, also more randomized. While 81% of the CDP units experienced growth over the period, only 40% of the match units

TABLE 3

3rd Grade CAT Scores (Unit Average) for Match Units

	1986	1987	1988	1989
A*	64	C 66 (+ 3)	L 72 (+10)	B 69 (+7; +5)
B	64	D 65 (+ 4)	C 68 (+ 2)	J 67 (+1; -12)
C	63	B 62 (- 2)	J 66 (+ 7)	E 67 (+5; +7)
D	61	L 62 (+13)	B 62 (N.C.)	C 66 (-2; +3)
E	60	J 59 (+ 4)	E 62 (+ 4)	A 57 (+2; -7)
F	59	A 58 (- 6)	D 58 (- 7)	H 57 (+3; NC)
G	58	E 58 (- 2)	M 57 (+15)	F 56 (NC; -3)
H	57	H 56 (- 1)	A 57 (- 1)	G 56 (+2; -2)
I*	56	G 55 (- 3)	K 56 (+ 1)	L 53 (-19; +4)
J*	55	K 55 (+ 1)	F 56 (+ 4)	D 53 (-5; -8)
K*	54	F 52 (- 7)	G 54 (- 1)	M 48 (-9; NC)
L	49	N 46 (+ 3)	H 54 (- 2)	K 47 (-9; -7)
M*	48	I 43 (-13)	I 45 (+ 2)	N 43 (-1; NC)
N*	43	M 42 (- 6)	N 44 (- 2)	I 42 (-3; -14)
O*	31	O 36 (+ 5)	O 35 (- 1)	O 42 (+7; +11)

* Units participated in TPAS Pilot (1985-87)

TABLE 4

Range and Decile Distributions of Match Units Per Year

	Range	Deciles : Units
1986	31 - 64 (33 points)	30th: 1; 40th: 3; 50th: 6; 60th: 5; 70th: 0
1987	36 - 66 (30 points)	30th: 1; 40th: 3; 50th: 7; 60th: 4; 70th: 0
1988	35 - 72 (37 points)	30th: 1; 40th: 2; 50th: 7; 60th: 4; 70th: 1
1989	42 - 69 (27 points)	30th: 0; 40th: 5; 50th: 6; 60th: 4; 70th: 0

experienced growth. Taken together, the match units show no clear patterns, except for some growth among the top four units, in 1988-89. Even here, however, the gain is randomized. Units B, E, and C had scores at least at the 60th percentile in 1985-86 and still do in 1988-89. However, Units A and D have dropped by -7 and -8 points respectively and have been replaced by Unit J, with a healthy 12 percentile point gain.

Grade 6 Results

Tables 5 and 6 present the same information for Grade 6 in the CDP pilot units. Generally speaking, achievement test scores for 6th grade students are lower than they are for 3rd grade students for several reasons, mostly having to do with the greater curricular freedom experienced in Grades 4, 5, and 6 and the increased likelihood that the test and the curriculum will diverge. We see this phenomenon clearly in Table 5 as compared to Table 1. However, Table 5 shows similar patterns of achievement in other respects.

In 6th grade for CDP units we see that the number of units ending below the 50th percentile increases by one, although with two exceptions, scores are up. It is noteworthy, that the floor rises from the 33rd percentile to the 37th. While the ceiling also rises (from the 69th to 75th), it is interesting that in 1986 only three units were above the 60th percentile, but this number increases to 5 in 1989, in addition to 1 above the 70th. Of the pilot units, 12 experienced improvements over the period, and four experienced declines of generally modest amounts, although the 5 point decline in Chowan is troublesome. The improvement scores range as high as +10 for one unit, with one at +7, two at +6, one at +5, three at +4, three at +3, and one at +1.

TABLE 5

6th Grade CAT Scores (Unit Average) for CDP Units

1986		1987		1988		1989	
Burlington	69	Burlington	72 (+ 3)	Burlington	71 (- 1)	Burlington	75 (+4; +6)
Haywood	63	Haywood	66 (+ 3)	N. Hanover	66 (+ 4)	Haywood	66 (+4; +3)
R. Rapids	62	R. Rapids	65 (+ 3)	Buncombe	63 (+ 2)	Buncombe	65 (+2; +7)
Burke	59	N. Hanover	62 (+ 4)	Burke	62 (+ 5)	Burke	65 (+3; +6)
N. Hanover	58	Buncombe	61 (+ 3)	Haywood	62 (- 4)	N. Hanover	61 (-5; +3)
Buncombe	58	Alexander	58 (+ 1)	Chowan	59 (+10)	Mecklenburg	60 (+4; +3)
Mecklenburg	57	Burke	57 (- 2)	Alexander	57 (- 1)	R. Rapids	59 (+3; -3)
Alexander	57	Mecklenburg	55 (- 2)	R. Rapids	56 (- 9)	Alexander	58 (+1; +1)
Chowan	54	Salisbury	54 (+16)	Mecklenburg	56 (+ 1)	Harnett	55 (+4; +5)
Orange	53	Orange	54 (+ 1)	Orange	52 (- 2)	Orange	52 (NC; -1)
Harnett	50	Perquimans	53 (+ 8)	Harnett	51 (+ 2)	Chowan	49 (-10; -5)
Tarboro	48	Montgomery	50 (+ 6)	Montgomery	50 (N.C.)	Perquimans	49 (-1; +4)
Perquimans	45	Tarboro	50 (+ 2)	Perquimans	50 (- 3)	Montgomery	48 (-2; +4)
Montgomery	44	Chowan	49 (- 5)	Tarboro	49 (- 1)	Salisbury	48 (+6; +10)
Salisbury	38	Harnett	49 (- 1)	Salisbury	42 (-12)	Tarboro	46 (-3; -2)
Greene	33	Greene	39 (+ 6)	Greene	42 (+ 3)	Greene	37 (-5; +4)

TABLE 6

Range and Decile Distributions of CDP Units Per Year

	Range	Deciles : Units
1986	33 - 69 (36 points)	30th: 2; 40th: 3; 50th: 8; 60th: 3; 70th: 0
1987	39 - 72 (33 points)	30th: 1; 40th: 2; 50th: 8; 60th: 4; 70th: 1
1988	42 - 71 (29 points)	30th: 0; 40th: 3; 50th: 8; 60th: 4; 70th: 1
1989	37 - 75 (38 points)	30th: 1; 40th: 5; 50th: 4; 60th: 5; 70th: 1

Again, as in the 3rd grade profile, we find a general pattern of improvement among the pilot units, with the bottom units generally moving up steadily and rapidly. No unit lost ground in both 3rd and 6th grade. However, nine units experienced improvements in both 3rd and 6th grades over the period 1985-89. When the data in Tables 7 and 8 are examined, these same patterns do not emerge.

Tables 7 and 8 present the achievement data for Grade 6 in the match units. The general level of achievement for this group starts out lower than for CDP units and improves modestly over the period. While the ceiling rises to the 65th percentile, the floor drops from the 37th percentile in 1986 to the 33nd in 1989. Moreover, six units fail to attain the median (50th percentile) in 1989, as compared to 8 in 1986, 7 in 1987, and 9 in 1988. Among individual units, eight experienced increases, two registered no change over the period, and five experienced declines, with four of these experiencing declines of 3 percentile points or more. Of the five units with the lowest scores in 1986, three are worse off in 1989. Moreover, only two units (L and G) begin the period in the 4th decile, and end the period above the median.

As Table 8 demonstrates, the distance between the lowest and highest scoring units increased over the period, caused by the decline in the low performers and some improvement at the top. If any pattern emerges from these data, it is that the poor performers are performing more poorly at the end of the period. Of eight units below the median in 1986, four lost ground by 1989 and only two rose to the national median of 50 or better. When the data in Table 7 are compared with those in Table 3, it becomes clear that Units B, C, E, and J experienced growth in both 3rd and 6th grade over the period. However, Units A, F, and I lost ground in both 3rd and 6th grades. Compare this with the nine units in the CDP group that improved both 3rd and 6th grade performance, while none experienced decline. Clearly,

TABLE 7

6rd Grade CAT Scores (Unit Average) for Match Units

1986		1987	1988	1989
A*	63	A 67 (+ 4)	A 63 (- 4)	C 65 (+3; +2)
C	63	C 62 (- 1)	C 62 (N.C.)	A 62 (-1; -1)
L	56	E 61 (+ 7)	E 62 (+ 1)	E 59 (-3; +5)
B	56	B 61 (+ 5)	B 58 (- 3)	B 59 (+1; +3)
J*	55	J 56 (+ 1)	J 54 (- 2)	J 56 (+2; +1)
E	54	L 56 (N.C.)	K 50 (+ 6)	K 56 (+6; +14)
G	52	G 55 (+ 3)	G 49 (- 6)	G 56 (+7; +4)
F	48	H 51 (+ 5)	L 49 (- 7)	L 56 (+7; NC)
D	48	F 47 (- 1)	H 48 (- 3)	H 54 (+6; +8)
M*	47	D 46 (- 2)	M 47 (+ 6)	M 49 (+2; +2)
H	46	K 44 (+ 2)	D 44 (- 2)	D 48 (+4; NC)
N*	43	M 41 (- 6)	F 42 (- 5)	F 45 (+3; -3)
K*	42	I 41 (+ 2)	O 41 (+10)	N 38 (NC; -5)
I*	39	N 39 (- 4)	N 38 (- 1)	I 36 (+4; -3)
O*	37	O 31 (- 6)	I 32 (- 9)	O 33 (-8; -4)

* IPAS Pilot Units 1985-87

TABLE 8

Range and Decile Distributions of Match Units Per Year

	Range	Deciles : Units
1986	37 - 63 (26 points)	30th: 2; 40th: 6; 50th: 5; 60th: 2; 70th: 0
1987	31 - 67 (36 points)	30th: 2; 40th: 5; 50th: 4; 60th: 4; 70th: 0
1988	32 - 63 (31 points)	30th: 2; 40th: 7; 50th: 3; 60th: 3; 70th: 0
1989	33 - 65 (32 points)	30th: 3; 40th: 3; 50th: 7; 60th: 2; 70th: 0

a very different pattern emerges whether the analysis is grade for grade, year for year, across the period, or for component units in the two groups.

Grade 8 Results

The data presented in Tables 9 and 10 illustrate 8th grade achievement in CDP pilot units. At first glance, they appear somewhat troublesome. For the first time, we see the range of difference in CDP units expanding, signifying that the ceiling is moving up faster than the floor or that the floor is actually dropping, or both. However, a second look that considers only 15 of the 16 units is reassuring. The same patterns of achievement seen earlier are confirmed. While the Greene County data cannot be ignored, they seem to reflect a pattern of only one. The number of units scoring below the 50th percentile in 1989 is only one quarter of those below the 50th in 1986. Moreover, of the four lowest scoring units in 1989, only Greene is actually losing ground and the other units are improving. Overall, thirteen units experienced score increases during the period, while two units declined. Seven units experienced increases of +4 or better. Thus, improvement was generalized among high, medium, and low scoring units over the period. This is in marked contrast to the performance of 8th grade students in the 15 match units. Tables 11 and 12 present the data for these students.

Table 12 shows that the group floor remained unchanged while the ceiling rose over the period. In general, then, achievement was mixed for the match group. Almost as many units failed to achieve at or above the national median in 1989 as in 1986, and only two of these low- scoring units marked improvements over the period. For the most part, units below the median in 1986 were still there in 1989, with the exception of Unit L. Four of the top five units experienced improvement over the period, while the middle units were mixed.

TABLE 9

8th Grade CAT Scores (Unit Average) for CDP Units

1986		1987		1988		1989	
R. Rapids	65	Burlington	64 (N.C.)	Burlington	68 (+ 4)	Burlington	67 (-1; +3)
Burlington	64	Haywood	62 (+ 1)	R. Rapids	65 (+ 4)	Haywood	63 (+1; +2)
Haywood	61	Buncombe	62 (+ 4)	Haywood	62 (N.C.)	Salisbury	63 (+9; +10)
Buncombe	58	R. Rapids	61 (- 4)	Chowan	57 (+10)	R. Rapids	62 (-3; -3)
Tarboro	54	N. Hanover	55 (+ 4)	N. Hanover	57 (+ 2)	Buncombe	62 (+6; +4)
Salisbury	53	Perquimans	53 (+ 2)	Buncombe	56 (- 6)	N. Hanover	58 (+1; +7)
Alexander	52	Montgomery	52 (+ 3)	Mecklenburg	55 (+ 4)	Alexander	56 (+1; +4)
Perquimans	51	Harnett	51 (+ 2)	Alexander	55 (+ 6)	Perquimans	56 (+5; +5)
N. Hanover	51	Mecklenburg	51 (N.C.)	Salisbury	54 (+ 9)	Mecklenburg	54 (-1; +3)
Mecklenburg	51	Burke	51 (+ 1)	Burke	53 (+ 2)	[illegible]	54 (+1; +5)
Chowan	50	Orange	50 (+ 1)	Harnett	53 (+ 2)	Tarboro	54 (+2; NC)
Burke	50	Alexander	49 (- 3)	Tarboro	52 (+ 3)	Burke	53 (NC; +3)
Montgomery	49	Tarboro	49 (- 5)	Orange	52 (+ 2)	Orange	53 (+1; +4)
Harnett	49	Chowan	47 (- 3)	Perquimans	51 (- 2)	Chowan	52 (-5; +2)
Orange	49	Salisbury	45 (- 8)	Montgomery	49 (- 3)	Montgomery	50 (+1; +1)
Greene	40	Greene	38 (- 2)	Greene	36 (- 2)	Greene	38 (+2; -2)

TABLE 10

Range and Decile Distributions of CDP Units Per Year

	Range	Deciles : Units
1986	40 - 65 (25 points)	30th: 0; 40th: 4; 50th: 9; 60th: 3; 70th: 0
1987	38 - 64 (26 points)	30th: 1; 40th: 4; 50th: 7; 60th: 4; 70th: 0
1988	36 - 68 (32 points)	30th: 1; 40th: 1; 50th: 11; 60th: 3; 70th: 0
1989	38 - 67 (29 points)	30th: 1; 40th: 0; 50th: 10; 60th: 5; 70th: 0

TABLE 11

8th Grade CAT Scores (Unit Average) for Match Units

1986		1987	1988	
C	59	B 58 (+ 3)	C 62 (+ 6)	E 63 (+ 8; + 9)
B	55	C 56 (- 3)	B 57 (- 1)	C 61 (- 1; + 2)
G	55	G 55 (N.C.)	E 55 (+ 1)	B 61 (+ 4; + 6)
E	54	K 55 (+ 6)	G 53 (- 2)	G 55 (+ 2; NC)
A	53	E 54 (N.C.)	A 53 (+ 4)	J 51 (+ 5; + 2)
J*	52	J 53 (+ 1)	K 51 (- 4)	A 52 (- 1; - 1)
H	51	L 52 (+ 3)	H 51 (+ 1)	H 52 (+ 1; + 1)
L	49	H 50 (- 1)	J 49 (- 4)	L 50 (+ 3; + 1)
K*	49	A 49 (- 4)	L 47 (- 5)	K 47 (- 4; - 2)
M*	45	D 48 (+ 3)	M 46 (+ 2)	D 44 (- 1; - 1)
F	45	M 44 (- 1)	O 46 (+ 9)	I 44 (NC; + 4)
D	45	N 41 (+ 1)	N 45 (+ 4)	M 42 (- 4; - 3)
I*	40	I 37 (- 3)	D 45 (- 3)	O 42 (- 4; + 5)
N*	40	F 37 (- 8)	F 44 (+ 7)	F 41 (- 3; - 4)
O*	37	O 37 (N.C.)	I 44 (+ 7)	N 37 (- 8; - 3)

* Units participated in TPAS pilot 1985-87

TABLE 12

Range and Decile Distributions of Match Units Per Year

	Range	Deciles : Units
1986	37 - 59 (22 points)	30th: 1; 40th: 7; 50th: 7; 60th: 0; 70th: 0
1987	37 - 58 (21 points)	30th: 3; 40th: 4; 50th: 8; 60th: 0; 70th: 0
1988	44 - 62 (18 points)	30th: 0; 40th: 8; 50th: 6; 60th: 1; 70th: 0
1989	37 - 63 (26 points)	30th: 1; 40th: 6; 50th: 5; 60th: 3; 70th: 0

Overall, these units did not do as well, relatively or absolutely, as did the CDP units. The match group had a slightly tighter range throughout the period as compared to the CDP units, partly because the match group's ceiling never equaled the CDP ceiling and because the match group's floor stayed steady. However, the individual units did not do as well as the individual CDP units. Only eight of the matched units made improvements over the period, while six lost ground.

Overall Results

An overview of the net changes over the period at each grade level for each unit is provided in Table 13. While six of the match units made gains in 3rd grade, 13 of the CDP units experienced improvements. At the 6th grade level, eight of the match units made gains, while 12 of the CDP units improved. Among the match units, eight experienced gains in 8th grade compared with 13 of the CDP units. Moreover, eight CDP units--Perquimans, New Hanover, Montgomery, Salisbury, Mecklenburg, Burke, Burlington, and Buncombe--made gains at each grade level, while no unit lost at all levels. In the match units, B, C, E, and J experienced improvements at all levels. However, Units A, and F experienced losses at each grade level. When it is remembered that the match units were selected because of their resemblance to the CDP units on critical characteristics, the markedly different patterns are difficult to reconcile, unless the Career Development Program effect is taken into account. While it would be foolhardy to attribute all of the differences to CDP, it would be equally foolish to contend that CDP made no difference.

Clearly, patterns of improved student performance have been demonstrated in the CDP units, both as a group and in most individual units. Moreover, the growth seems as likely to have occurred among high-performing units as in middling or low-performing units. This improvement appears to be as likely in

TABLE 13

Overview of Change in 15 Match Units and 16 CDP Units

Unit	Grade 3	6	8
A	-	-	-
B	+	+	+
C	+	+	+
D	-	NC	-
E	+	+	+
F	-	-	-
G	-	+	NC
H	NC	+	+
I	-	-	+
J	+	+	+
K	-	+	-
L	+	NC	+
M	NC	+	-
N	NC	-	-
O	+	-	+
Total	6+	8+	8+
	7-	5-	6-
	3NC	2NC	1NC

Unit	Grade 3	6	8
Perquimans	+	+	+
Chowan	+	-	+
N. Hanover	+	+	+
Greene	+	+	-
Harnett	-	+	+
R. Rapids	+	-	-
Tarboro	+	-	NC
Montgomery	+	+	+
Salisbury	+	+	+
Mecklenburg	+	+	+
Alexander	-	+	+
Burke	+	+	+
Orange	+	-	+
Burlington	+	+	+
Haywood	-	+	+
Buncombe	+	+	+
Total	13+	12+	13+
	3-	4-	2-
	ONC	ONC	1NC

\+ = improved in 1989 over 1986

\- = declined in 1989 over 1986

NC = No change in 1989 over 1986

3rd or 6th grade as 8th grade. By contrast, we see no clearly defined patterns among the individual match units, although it could be argued that, as a group, the match units have experienced some improvement. This, of course, would reflect the general statewide improvement in CAT scores. This generalized improvement, however, is quite inconsistently demonstrated in the individual units, in contrast with individual CDP units.

Student to Student Comparison

Before concluding the presentation of the CDP unit data in contrast to the match unit data, one more phenomenon should be discussed. In 1963, John B. Carroll published one of the most influential articles in the educational literature. (Carroll, 1963). In "A Model of School Learning", Carroll posited that students would learn as a function of aptitude, perseverance, opportunity, and quality of instruction. While no one of the these factors could overcome the others, and the interaction was not clear, obviously students will differ in aptitude and perseverance. Opportunity and quality of instruction, however, are more or less controllable by the institution. We would assume, then, that score changes in the same population, over time, could reflect a change in opportunity to learn and/or quality of instruction. That is, if we compared scores of a group of students over time, some part of the difference should be attributable to quality of instruction. Tables 14-16 provides us with the opportunity to test the hypothesis with the units we have been considering. Table 14 presents the achievement data for 6th grade students in 1986 and the data for the same students two years later, that is, 8th grade in 1988 for both groups of units. Table 15 compares performance of 1986-87 6th graders to themselves in 1988-89 when they were in Grade 8. Table 16 shows the scores for 1985-86 3rd graders, compared to the same group in 1988-89, when they were in 6th grade.

TABLE 14

Comparison of CAT Scores for 6th Grade (86) and 8th Grade (88) for CDP and Match Units

UNIT	6th Gr(86)	8th Gr(88)	Diff	UNIT	6th Gr(86)	8th Gr(88)	Diff
Burlington	69	68	-1	C	63	62	-1
Haywood	63	62	-1	A	63	53	-10
Roanoke Rapids	62	65	+3	L	56	47*	-9
Burke	59	53	-6	B	56	57	+1
N. Hanover	58	57	-1	J	55	49*	-6
Buncombe	58	56	-2	E	54	55	+1
Mecklenburg	57	55	-2	G	52	53	+1
Alexander	57	55	-2	F	48*	44*	-4
Chowan	54	57	+3	D	48*	45*	-3
Orange	53	52	-1	M	47*	46*	-1
Harnett	50	53	+3	H	46*	51	+5
Tarboro	48*	52	+4	N	43*	45*	+2
Perquimans	45*	51	+6	K	42*	51	+9
Montgomery	44*	49*	+5	I	39*	44*	+5
Salisbury	38*	54	+16	O	37*	46*	+9
Greene	33*	36*	+3				

* = Unit average below 50th percentile

TABLE 15
Comparison of CAT Scores for 6th Grade (87) and 8th Grade (89) for CDP and Match Units

	6th(87)	8th(89)		6th(87)	8th(89)
Burlington	72	67 (- 5)	A	67	52 (-15)
Haywood	66	63 (- 3)	C	62	61 (- 1)
R. Rapids	65	62 (- 3)	E	61	63 (+ 2)
N. Hanover	62	58 (- 4)	B	61	61 (NC)
Buncombe	61	62 (+ 1)	J	56	54 (- 2)
Alexander	58	56 (- 2)	L	56	50 (- 6)
Burke	57	53 (- 4)	G	55	55 (NC)
Mecklenburg	55	54 (- 1)	H	51	52 (+ 1)
Salisbury	54	63 (+ 9)	F	47	41 (- 6)
Orange	54	53 (- 1)	D	46	44 (- 2)
Perquimans	53	56 (+ 3)	K	44	47 (+ 3)
Montgomery	50	50 (NC)	M	41	42 (+ 1)
Tarboro	50	54 (+ 4)	I	41	44 (+ 3)
Chowan	49	52 (+ 3)	N	39	37 (- 2)
Harnett	49	54 (+ 5)	O	31	42 (+ 9)
Greene	39	39 (NC)			

Number of Units Gaining/Losing Percentile Points

CDP

					1	2	2	1	2	2	1		2	1	1				1	
-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
1				2				3	1	2	2	1	2						1	

Match

TABLE 16
Comparison of CAT Scores for 3rd Grade (86) and 6th Grade (89) for CDP and Match Units

	3rd(86)	6th(89)		3rd(87)	6th(89)
Burlington	76	75 (- 1)	A	64	62 (- 2)
Haywood	71	66 (- 5)	B	64	59 (- 5)
R. Rapids	67	59 (- 8)	C	63	65 (+ 2)
Buncombe	67	65 (-12)	D	61	48 (-13)
Harnett	64	55 (+ 9)	E	60	59 (- 1)
Burke	64	65 (+ 1)	F	59	45 (-14)
N. Hanover	61	61 (NC)	G	58	56 (- 2)
Tarboro	61	46 (-15)	H	57	54 (- 3)
Mecklenburg	60	60 (NC)	I	56	36 (-20)
Orange	59	52 (- 7)	J	55	56 (+ 1)
Alexander	59	58 (- 1)	K	54	56 (+ 2)
Montgomery	52	48 (- 4)	L	49	56 (+ 7)
Perquimans	48	49 (+ 1)	M	48	49 (+ 1)
Greene	47	37 (-10)	N	43	38 (- 5)
Salisbury	44	48 (+ 4)	O	31	33 (+ 2)
Chowan	43	49 (+ 6)			

Number of Units Gaining/Losing Percentile Points

CDP

-10	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	10+
2	1	1	1	1			1	1			2	2	2			1		1				

Match

-10	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	10+
3						2		1	2	1		2	3					1				

First, let us consider only the CDP units. In 1986, five units (indicated by asterisk) had CAT averages below the 50th percentile. In 1988, only two units were below the 50th, and one of these was at 49th. Students in three units, then, had improved their scores enough to propel themselves over this line of demarcation. The neediest students, in short, performed better. Unfortunately, this pattern of improvement was <u>not</u> universal in the CDP units. Eight units declined between 6th and 8th grade. However, these losses were, by and large modest. Four units lost one percentile point, while three lost two points. One unit lost six points.

However, among the eight gaining units, the gains were significant. The <u>smallest</u> gain was three percentile points (experienced by four units), while other gains of four, five, six, and sixteen points were gained:

Points	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+16
Units	1	0	0	0	3	4	0	0	0	4	1	1	1	1

In short, the gainers improved more than the losers lost. Moreover, among the lower half of the units (arranged by 6th grade scores) only <u>one</u> lost ground. Children who had done well in 6th grade continued to do well and children who had done relatively poorly in 6th grade made remarkable growth.

Now let us turn to the right side of the table. The picture here is somewhat different. Notice that in 1986, eight units (F, D, M, H, N, K, I, and O) all scored below the 50th percentile, in Grade 6. In 1988, eight units were still below the 50th percentile. However, only Units H and K had moved

from below to above the 50th percentile. Units F, D, M, N, I, and O were still below the median, although three of them (N, I, and O) had improved. Unfortunately, F, D, and M had actually declined further. More unfortunately, Units L and J had also fallen below the median, going from 56th and 55th to 47th and 49th respectively. While the very neediest students had improved their scores, in some cases dramatically, many of the "near needy" had declined markedly. The trend for marginal units (those near, but below the median) was clearly down from the median.

Moreover, among the seven top-scoring units in 1988, (the only ones above the median in 1986), four had lost ground as compared to themselves when in Grade 6. The losses were not small. While the highest scoring district had lost only 1 point, the other units lost 10, 9, and 6 percentile points. The only gains in the top seven were one percentile point each in three units:

Points	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
Units	1	1	0	0	1	0	1	1	0	2	0	3	1	0	0	2	0	0	0	2	0

It must not be overlooked that the five lowest-scoring 6 grades all made improvement, repeating the pattern seen among CDP's lowest scoring units. Their growth, however, was not as large as the growth of the CDP units, while the losing units tended to lose more than the losing CDP units. Clearly, the comparison favors CDP, indicating that sustained growth occurred in the majority of these units with relatively modest losses.

Interestingly, in a similar comparison for another group of students, the same patterns appear. Table 15 compares students' performance for 1987 6th graders with their performance as 8th graders in 1988-89. Once again, we observe that, for CDP units, the number of districts below national median

declined from three in 1986-87 to one in 1988-89. In the match units, the number below the median stayed constant at 7 for both grades.

Looking only at CDP units, we note that seven of the eight highest-achieving sixth grades lost one to five percentile points in the eighth grade comparison. Among the eight lower performing sixth grades, however, we find higher percentile scores at eighth grade for five units, with two experiencing no change. The only one of this sub-group that lost, experienced only a one percentile point change. Among these gaining units, increases ranged from three to nine percentile points.

In the match unit group, the patterns is weaker than we saw in Table 14. Among the high-performing 6 grades, two remained unchanged in Grade 8, two had increases of one or two percentile points and four had losses of one to 15 points. While this group included fewer losing units than was true for CDP, the losses were somewhat larger. Among the lower performing units, the pattern of improvement demonstrated in Table 14 does not continue, however. While four of these units did make improvements in Grade 8, three ended up with lower scores.

Table 16 provides data related to changes for students who were in Grade 3 in 1985-86 and Grade 6 in 1988-89. It is important to realize that, generally speaking, students at Grade 6 test lower than students at Grade 3 for a variety of reasons. As Table 16 shows, students in CDP 6 grades did not perform as well as they had in 3rd grade. However, among the four 3rd grades below the national norm in 1985-86, three showed improvements at the 6 grade, although the 10 percentile point loss in Greene County is serious. Although the numbers are larger, the data in Table 16 replicate th same patterns we saw in Tables 14 and 15.

Considering all of the data presented in Tables 1-16, what conclusions may be drawn? These data show that:

- Among CDP units, in Grades 3 and 8, the number scoring below the 50th percentile declined over the period 1986-89. In 1989, only one unit was below the median in Grade 3, six in Grade 6, and one in Grade 8. By contrast, five of the match units were below the 50th percentile in Grade 3, up from four. Six units were below the median in 6th grade in 1989, down from eight. In 8th grade seven units were below the national average, down from eight.
- In Grades 3 and 6, CDP units raised the floor from the 43rd and 33rd percentiles to the 49th and 37th percentile, respectively. While the floor for 8th grade fell from the 40th to the 38th percentile, if we discount a single unit, the floor rose slightly to the 50th percentile over the period. In the match units, the floor rose in Grades 3 and 6, but fell in Grade 8.
- Between 75% and 81% of the CDP units experienced growth in each grade over the period. Between 40% and 53% of the match units improved at each grade level.
- Decreases of scores over the period were much more likely to occur among units in the match set.
- If we compare 6th grades to 8th grades, we see that units of the match set were more likely to lose, and to lose more, than units in the CDP set.
- Finally, no CDP unit experienced declines in all three grades over the period. In the match units, two units experienced losses in every grade. Among the CDP set, eight units made gains in each grade. In the match set, only four units equalled this achievement.

CONCLUSION

It is clear from the preceding analysis that, as Larry Lazotte has said, school units are either improving or declining. When student achievement in the Career Development Program pilot units is examined, the general tendency is toward improvement. Fewer CDP units scored below the national median for 3 grade in 1989 than did in 1986. The same was true for 8th grade achievement. In the group of match units, we did see a decline of units scoring below the median in Grade 3, but we saw no improvement on this parameter in Grades 6 or 8.

In general, CDP units, for the most part, experienced improvement over the period in Grades 3, 6, and 8. In Grade 3, 13 CDP units improved. In Grade 6, 12 of the CDP units improved or showed no change. In Grade 8, 13 CDP units improved. The corresponding numbers for the match group were six units for Grade 3, eight units for Grade 6, and eight for Grade 8.

As Table 13 shows, improvement was more likely to occur in the CDP units than in the matched set. If we think of each grade per unit as a measurement opportunity, the CDP units as a group improved on 79% of the measures, stayed unchanged on 2%, and declined on 19% of the measures. The matches' corresponding rates were 49% improvements, 13% no change, and 38% decline. Whereas clear patterns of improvement are demonstrated by the CDP units, we would characterize the change in the match units as random improvement or diffused change. While it is unreasonable to credit all the improvement experienced in the CDP units to their efforts in the pilot, certainly some of the improvement should be credited to the pilot. It is important to remember that these analyses were undertaken precisely because of the almost universal sentiment among CDP participants that instruction had improved <u>because</u> of the unit's participation in the pilot.

Perhaps the most interesting demonstration of CDP's efficacy is presented in Tables 14 and 15. Remember that these scores were posted by largely the same population at the beginning of the CDP experience and two years later. The direction, location, and amount of improvement are heartening. All of the units scoring below the national median in 1986 showed sizeable improvements. While decreases in the high- and middle-scoring units is cause for concern, the relatively small decreases, with one exception, do not obviate the benefits of the CDP.

By contrast, in the match units, while some of the lowest performing units do make commendable gains, the relatively larger, more wide-spread losses among the other units are noteworthy. Again, the random nature of the changes offers confirmation that quite different conditions obtained in this group from the CDP group.

It has been seriously suggested that students would suffer because of a perceived necessity for "robotized" teaching in the CDP units. These data put the lie to that charge. Clearly, students in the CDPs, taken as a group, have benefitted from more wide-spread good teaching than did students in the match sample. While we cannot factor out how much of the achievement gains to credit with specific features of Career Development or other innovations, for that matter, it is clear that the 16 participating units have posted significant gains in student achievement, in both relative and absolute terms.

REFERENCES

Brandt, Richard; Duke, Daniel; French, Russell; and Iwanicki, Edward. 1988. "A Review With Recommendations of the N.C. Teacher Performance Appraisal Instrument." Raleigh, N. C.: NC General Assembly.

Brophy, Jere. 1980. "Recent Research on Teaching". East Lansing, MI: Institute for Research on Teaching, Michigan State University.

Brophy, Jere. 1988. "Research on Teacher Efects: Uses and Abuses." Elementary School Journal. Vol. 89, #1, pp. 3-21.

Bunche, Michael B. 1988. Career Development Plan Pilot Project Survey: Final Report. Durham, N. C.: Measurement, Inc.

Carroll, John G. 1963. "A Model of School Learning". Teacher's College Record. New York, N. Y.: Columbia University.

Division of Personnel Relations, North Carolina Department of Public Instruction. 1986 Getting Ready: A Report of Teachers' and Administrators' Belief About the Implementation of the North Carolina Career Development Program. Raleigh. N. C.: North Carolina Department of Public Instruction.

Division of Research, North Carolina Department of Public Instruction. 1986. Report of Student Performance. Raleigh, N. C.: North Carolina Department of Public Instruction.

Division of Testing, North Carolina Department of Public Instruction. 1988. Report of Student Performance. Raleigh, N. C.: North Carolina Department of Public Instruction.

Gage, N.L. 1984. "What Do We Know About Teaching Effectiveness?" Phi Delta Kappan, October, pp. 87-93.

Gage, N.L. and Nudels, Margaret C. 1989. "Process-Product Research on Teaching: A Review of Criticisms." Elementary School Journal. Vol. 89, #3, pp. 253-300.

Holdzkom, David. 1987 "Teacher Appraisal In North Carolina". Educational Leadership. Vol. 44, #7, pp. 40-44.

Holdzkom, David; Kuligowski, Barbara; and Stacey, Dennis. 1989. "A Longitudinal Study of Teacher Performance in N.C's Career Development Program." Paper presented at the Annual Conference of AERA, San Francisco, CA.

North Carolina Association of Educators. 1986. N. C. Career Ladder Pilot Survey Analysis: A Summary of Major Findings. Raleigh, N. C.: NCAE.

Research and Testing Services, North Carolina Department of Public Instruction. 1987. Report of Student Performance. Raleigh, N. C.: North Carolina Department of Public Instruction.

Research and Service Institute. 1989. A Third-Party Evaluation of the North Carolina Career Development Program Pilot. Raleigh, N.C.: N.C. General Assembly.

Rosenshine, Barak. 1982. "Teaching Functions in Instructional Programs." Paper presented at the National Invitational Conference. "Research on Teaching: Implications for Practice". Sponsored by NIE.

Stacey, Dennis; Holdzkom, David; and Kuligowski, Barbara. 1989. "Effectiveness of the N.C. Teacher Performance Appraisal System." Journal of Personnel Evaluation in Education. Vol. 3, pp. 79-106.

Teddlie, Charles; Kirby, Peggy C.; and Stringfield, Sam. 1989. "Effective versus Ineffective Schools: Observable Differences in the Classroom." American Journal of Education. May, pp. 221-236.