DOCUMENT RESUME

ED 266 424 CS 008 319

AUTHOR Linn, Robert L.; Meyer, Linda A.
TITLE Kindergarten Instruction and Early Reading Achievement. Technical Report No. L-4.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Dec 85
CONTRACT 400-81-0030
NOTE 27p.; Paper presented at the Annual Meeting of the National Reading Conference (34th, St. Petersburg, FL, November 28-December 1, 1984). Figure 3 contains small print.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Feedback; Kindergarten; Primary Education; *Reading Achievement; *Reading Instruction; *Reading Research; Reading Skills; Reading Strategies; Teacher Behavior; *Teacher Effectiveness; Teacher Influence; Teaching Methods; Teaching Styles; *Time on Task

ABSTRACT

The relationship of the amount of classroom time devoted to reading instruction, the number of reading related activities, and teacher instruction feedback to reading achievement at the end of kindergarten was investigated for a sample of approximately 300 children in 14 kindergarten classrooms at three schools. Based on nine rounds of full-day observations, it was found that there are great between- and within-class differences in the amount and type of reading instruction received by the kindergarten children. These differences were strongly related to student decoding ability in the spring after controlling for fall achievement. Future analyses of the continuing longitudinal follow-up of these children will investigate the degree to which these differences in early reading achievement are reflected in later reading comprehension differences. (Tables of findings are included.) (Author/EL)

ED266424

CENTER FOR THE STUDY OF READING

Technical Report No. 384

KINDERGARTEN INSTRUCTION AND EARLY READING ACHIEVEMENT

Robert L. Linn
Linda A. Meyer
University of Illinois at Urbana-Champaign

(Longitudinal Study Report 4)

December 1985

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

The work upon which this publication is based was performed pursuant to Contract No. 400-81-0030 of the National Institute of Education. It does not, however, necessarily reflect the views of this agency. Paper presented at the 34th Annual Meeting of the National Reading Conference, November 29, 1984.

Abstract

The relationship of the amount of classroom time devoted to reading instruction, the number of reading related activities, and teacher instruction feedback to reading achievement at the end of kindergarten was investigated for a sample of approximately 300 children in 14 kindergarten classrooms at three schools. Based on nine rounds of full-day observations, it was found that there are large between- and within-class differences in the amount and type of reading instruction received by the kindergarten children. These differences were strongly related to student decoding ability in the spring after controlling for fall achievement. Future analyses of the continuing longitudinal follow-up of these children will investigate the degree to which these differences in early reading achievement are reflected in later reading comprehension differences.

Kindergarten Instruction and Early Reading Achievement

The purpose of this report is to present some of the results from the analyses of a large observational data set showing the relationships of kindergarten teacher behavior with student achievement in reading. The data were obtained as part of an ongoing longitudinal study (Meyer, Linn, & Hastings, 1985). The focus in this report is on just a part of the general model described by Meyer et al. (1985). Included in the analyses are fall and spring student test scores and classroom observational variables. Excluded from the analyses reported here are measures of instructional materials and family background variables.

The analyses reported are based on results for approximately 300 children in 14 kindergarten classes at three schools. A battery of tests measuring general verbal ability, listening skills, the recognition of letters and words, language production, and recall was administered at the beginning of kindergarten in the fall of 1983. A similar battery with the addition of reading tests and a test of science knowledge was administered in the spring of 1984. The teacher variables were based on nine rounds of classroom observations during the year. Each round consisted of a full day of observation. Information about what teachers were doing and how long they were doing it was recorded throughout the day. Although the teacher is the focus, interactions were tagged to particular children or to small groups of children as well as to the whole class as appropriate. When the class was divided into small groups or into "activity time" as it's called in one district, the observation procedure shifted to sweeping the classroom every 5 minutes to record where each child was and to count how many children were on task. Thus, there are two general types of variables that are derived from the observational data: those that are linked to individual children within a classroom and those that are linked only to the teacher or the classroom as a whole.

The model for the analyses is depicted in Figure 1. The total group relationships of student achievement in the spring with fall pretest performance and classroom observational variables are broken down into between-class and within-class components. The sample size for the between-class analyses is only 14. With this small number of cases, the number of variables that can usefully be considered is quite limited and even at that, the analyses have relatively low power. Due to the instability of relationships in the between-class analyses, the results must be interpreted with caution. Data for this same group of kindergarten teachers on a second cohort of children were collected in 1984-85, however, and the data for the second cohort will be used to replicate the findings obtained for the first cohort.

Insert Figure 1 about here.

Despite the limitations on the between-class analyses due to sample size, it is important to separate those effects from the within-class effects. Total-group relationships represent a confounding of the between- and within-class relationships and for that reason can be quite misleading. The within-class relationships depend on the differential instruction and interactions between the teacher and specific children or small groups of children. Homogeneous versus heterogeneous ability grouping within a class and differences in materials covered and in instructional time for different groups may affect the within-class relationships. Such effects are apt to be masked in total-group analyses, however, due to a confounding with between-class effects.

Previous research (e.g., Barrett, 1965; Durkin, 1974-75; Dykstra, 1967; Stevenson, Parker, Wilkinson, Hegion, & Fish, 1976) suggests that pre-kindergarten tasks involving the recognition of letters and words are among the most powerful early predictors of subsequent reading achievement. Our preliminary analyses yielded similar results. The Wide Range Achievement Test (WRAT) was used in the fall to assess the ability of children to recognize letters and words. The WRAT was found to have higher correlations with performance on the spring tests than any of the other measures in the fall battery. Because of its high correlations with the tests administered in the spring and because of the need to keep this presentation of results to a manageable size, we have chosen to use the WRAT as the primary control variable in the analyses reported here.

Test Score Means and Correlations

The fall classroom and school means and standard deviations on the WRAT are shown in Table 1. The means for the three schools are quite similar, ranging from 18.3 to 19.3. The differences between schools in overall means on the fall administration of the WRAT are not statistically significant. The classroom means range from 15.6 to 22.0, with the highest and lowest means occurring in the same school. The between-class differences in means are statistically significant only at school 3. The between-class differences in school 3 support the belief that there is some attempt to form homogeneous classes with regard to ability at that school. There is no such indication at schools 1 and 2, however.

Insert Table 1 about here.

The three spring tests of primary interest for this report are the WRAT, the Woodcock, and the Chicago. The WRAT is a retest and provides an indication of growth in ability to recognize letters and words between fall and spring of the kindergarten year. The Woodcock Reading Mastery Test is an individually administered test of a child's ability to identify words. For our sample of kindergarten children the test is relatively difficult and is of interest at this stage as a means of differentiating between children at the upper end of the distribution on several of the other tests in the battery. The Chicago Reading Test is a measure of a child's ability to decode. It was expected to be particularly sensitive to between-class differences in emphasis on decoding skills.

The classroom and school means and standard deviations on the three spring tests are listed in Table 2. Although the primary interest is in differences after adjusting for fall test scores, it is worth noting that the school means are significantly different on all three of these tests. The between school differences are most notable on the Chicago Test. On that test, all four classroom means in school 1 are markedly higher than any of the other classroom means. As was indicated by Meyer, Linn, Mayberry, and Hastings (1985), school 1 had the most highly structured and substantial reading curriculum among the three schools. The teachers devoted more time to phonics concepts and vocabulary than did any of the other seven teachers. The differences between classes within a school are statistically significant only on the Chicago at school 3, where the observational data indicated that there was the least consistency in the instructional practices of the teachers.

Insert Table 2 about here.

Listed in Table 3 are three types of correlations between the fall WRAT scores and scores on the spring tests. These are the total-group correlation, the between-class correlation, and the median, minimum, and maximum values of the within-class correlations. As can be seen, the total-group correlations are generally higher than either the between-class or the median within-class correlations. This reflects the confounding of between- and within-class relationships in the total-group correlations.

Insert Table 3 about here.

Between-Class Results

Given the small number of classrooms, separate between-class analyses were conducted for each of several classroom observational variables. Classroom means on the spring tests were regressed on the mean fall WRAT score and a single classroom observational variable. The unique contribution of an observational variable was assessed by the difference between the squared multiple correlation when the fall WRAT and the observational variable were used as predictors and the squared correlation of the fall WRAT alone. Results from these analyses are reported in Table 4.

Insert Table 4 about here.

The first row of numbers in Table 4 lists the squared multiple correlation between the classroom means on the fall WRAT and the classroom means on the three spring tests. The next two rows report the squared multiple correlation of the fall WRAT and the total number of reading related activities observed during the nine rounds of observations (RACT) and the difference between the squared multiple and the simple squared correlation for the fall WRAT alone. The later figure is an estimate of the unique contribution of the amount of reading related activities to average spring achievement after controlling for between-class differences in initial performance on the WRAT. The following pairs of rows in Table 4 report analogous results for the total time devoted to decoding instruction (D-Time) and the total amount of positive feedback given to students during reading related activities (POS-F).

With one exception, the increases in the squared multiple correlations are all statistically significant despite the small number of classes available for the analyses. The single exception is for total reading related activities when the Woodcock is used as the dependent variable. As might be expected from the distribution of classroom means on the spring tests that were reported in Table 2, the significant differences are due primarily to the fact that the teachers in school 1 spent more time in decoding instruction, had more reading related activities, and gave more positive feedback than teachers at the other two schools. This is so despite the fact that school 1 has a half-day kindergarten, whereas school 3 has a full-day kindergarten.

The observational variables provide fairly clear distinctions among teachers. Although the amount of time or the number of reading related activities for a given teacher varied from round to round, teachers with both an AM and a PM class were generally quite consistent across the two classes on a given day. Furthermore, the patterns across observational rounds are quite distinct for different teachers. These characteristics are illustrated in Figure 2 for one teacher from school 1 (denoted by the squares) and one teacher from school 2 (denoted by the triangles). As can be seen, the number of minutes devoted to instruction by the teacher from school 1 is consistently higher than the corresponding values for the teacher from school 2 from the third round on. Both teachers are relatively consistent in the amount of time devoted to instruction for their morning and afternoon classes. The teacher from school 2 devotes about half of the available 150 minutes per day to instruction. In contrast, starting in round 3, the teacher from school 1 devotes about two-thirds of the available time to instruction each day. Consistent differences such as these contribute to the significant differences in spring achievement that were reported in Table 4.

Insert Figure 2 about here.

Within-class Results

Regression analyses were used to investigate the within-class relationships. The initial analyses used effect codes for classes, the fall WRAT scores, and the interaction of the WRAT with the effect code variables. This analysis provided a test of the homogeneity of the within-class slopes for the regression of each spring test on the fall WRAT. The analyses also provide an upper boundary on the amount of variability in spring achievement that can be explained by initial achievement differences between students on the WRAT and any class level observational variables.

The proportions of spring test score variance that is predictable from the fall WRAT, knowledge of the classroom that a student is in, and the interaction of the fall WRAT with classroom (i.e., heterogeneity in the within-class slopes) are reported in Table 5 for the three spring tests. Also shown are the increases in the squared multiple correlations when the classroom effect code variables are added to the fall WRAT (unique-class) and when the interactions are added (unique-interaction). The latter value provides a test of the homogeneity of the within-class slopes.

Insert Table 5 about here.

As can be seen in Table 5, the interaction terms add significantly to the prediction of each of the spring tests, indicating that the within-class slopes are heterogeneous. If the slopes were not heterogeneous, the increases in the squared multiples due to the classroom effect code variables would provide an alternative test of the between-class differences based on the logic of analysis of covariance. However, since the within-class slopes are heterogeneous, it is important to investigate the differences in the slopes and possible reasons for these differences.

Figure 3 illustrates the within-class slopes for the spring scores on the Chicago Test. The within-class regressions for the four classes in school 1 are shown in the upper left-hand corner of the figure with squares used for teacher 1 and triangles for teacher 2. The slopes for the three teachers in school 3 are shown in the upper right-hand corner, while those for the first two teachers (4 classes) in school 2 are in the lower left-hand corner and the third and fourth teachers in school 2 in the lower right-hand corner.

Three features of Figure 3 are worthy of comment. First, the regressions for teachers with an AM and a PM class are quite similar. Second, the regressions for classes in school 1 are consistently higher than those for the classes in the other two schools. This reflects the previously discussed finding that the spring achievement is higher for the classes in school 1 than in the other two schools. The within-class regressions demonstrate that this difference holds across different levels of initial achievement on the WRAT. Third, the largest source of heterogeneity in the within-class slopes is due to school 3. It is worth recalling in this regard that school 3 is the one with the least uniformity in the curriculum. Without a specified reading curriculum, teachers in school 3 use quite different approaches and this variability seems to be reflected by the great variability in the level and slopes of the within-class regressions for the three classes in school 3.

Insert Figure 3 about here.

Given the variability in the within-class regressions and the relative consistency between the AM and PM classes for a given teacher, we decided to focus our analyses of the within-class relationships with teacher observational variables at the level of the teacher rather than the class. Consequently, effect code variables for teacher rather than class were used. The observational variables were disaggregated to the level of a student within a classroom. That is, the information that was linked to individual students or small groups of students was used as a student level variable.

Since the increments in the squared multiple correlations due to class effects codes and interactions combined were larger for the Chicago Test than for the other two spring tests, the final analyses reported here are limited to the Chicago Test. The squared multiple correlations and changes in squared multiples (reported in Table 6) are based on the student level observational variables. As can be seen in Table 6, roughly half the variance in the Chicago Test scores can be predicted from a combination of the fall WRAT scores and either the total reading related interactions with an individual student or the total amount of decoding instructional time received by an individual student. The addition of the former variable increases the proportion of variance predictable from the WRAT alone by .13, while the increase due to the latter is .14.

Insert Table 6 about here.

Although the addition of the eight effect code variables for teachers increases the proportion of variance slightly (by either .03 or .04), the majority of the possible increase can be explained by the number of reading related interactions or the total amount of decoding instructional time received by a student.

Discussion

The analyses presented in this report represent only a beginning of those to be done. We have yet to combine the observational and student achievement variables with other variables depicted in the general model described elsewhere (Meyer, Linn, & Hastings, 1985) and to begin to construct models that pull apart the underlying structure and the measurement errors associated with the variables. We will also be more confident about the stability of our findings when they have been replicated using data from the second cohort that are currently being collected. Finally, the real contribution of a longitudinal study such as this must await the collection and analysis of data from subsequent years. Nonetheless, we find these preliminary results encouraging.

There are clearly some sizeable differences in the instruction received by students in the three schools in this study. These differences seem to be reflected in achievement in the spring. Both between- and within-class differences in the amount of time devoted to decoding instruction and in the frequency of the interactions are related to student achievement after controlling for initial differences on the fall WRAT. It remains to be seen whether these initial differences will be reflected in future differences in reading comprehension.

References

Barrett, T. C. (1965). Visual discrimination tasks as predictors of first-grade reading achievement. Reading Teacher, 18, 276-282.

Durkin, D. (1974-75). A six year study of children who learned to read in school at the age of four. Reading Research Quarterly, 1, 9-61.

Dykstra, R. (1967). The use of reading readiness tests for prediction and diagnosis: A critique. In T. C. Barrett (Ed.), The evaluation of children's reading achievement. Newark, DE: International Reading Association.

Meyer, L. A., Linn, R. L., & Hastings, C. N. (1985). The heuristic and measurement models driving a study of the development of reading comprehension and science knowledge (Tech. Rep. No. L-2). Urbana: University of Illinois, Center for the Study of Reading.

Meyer, L. A., Linn, R. L., Mayberry, P. W., & Hastings, C. N. (1985). A look at instruction in kindergarten: Observations of interactions in three school districts (Tech. Rep. No. L-3). Urbana: University of Illinois, Center for the Study of Reading.

Stevenson, H. W., Parker, J., Wilkinson, A., Hegion, A., & Fish, E. (1976). Longitudinal study of individual differences in cognitive development and scholastic achievement. Journal of Educational Psychology, 68(4), 377-400.

Table 1

Fall 1983 Classroom Means and Standard Deviations on the Wide Range Achievement Test (WRAT)

Teacher		Mean	Standard Deviation
		School 1	
Teacher 1:	AM	17.1	6.60
	PM	19.1	8.53
Teacher 2:	AM	20.6	5.84
	PM	20.5	9.45
Total:		19.3	7.69
		School 2	
Teacher 1:	AM	19.5	6.42
	PM	18.5	8.80
Teacher 2:	AM	16.1	6.58
	PM	19.3	6.06
Teacher 3:	AM	21.3	4.86
	PM	18.6	7.69
Teacher 4:	AM	20.0	6.64
Total:		19.1	6.82
		School 3	
Teacher 1		15.6	8.57
Teacher 2		22.0	6.12
Teacher 3		18.3	7.57
Total:		18.6	7.83

Table 2

Spring 1984 Classroom and School Means and Standard Deviations (in parentheses) for Three Tests

		WRAT	Woodcock	Chicago
		School 1		
Teacher 1:	AM	28.2 (4.43)	1.6 (3.11)	44.1 (21.94)
	PM	29.9 (5.27)	4.3 (4.61)	52.4 (26.50)
Teacher 2:	AM	30.4 (9.41)	3.5 (4.90)	48.1 (24.04)
	PM	30.3 (7.68)	5.2 (5.97)	49.6 (30.21)
Total:		29.7 (6.94)	3.6 (4.85)	48.4 (25.48)
		School 2		
Teacher 1:	AM	28.5 (5.77)	2.1 (2.83)	27.9 (21.97)
	PM	28.6 (5.99)	1.8 (3.33)	23.0 (23.05)
Teacher 2:	AM	25.4 (3.62)	.8 (1.35)	21.4 (11.77)
	PM	26.4 (2.96)	2.6 (3.08)	24.8 (18.23)
Teacher 3:	AM	27.0 (2.66)	1.6 (2.30)	21.4 (15.07)
	PM	27.5 (6.29)	1.9 (2.82)	26.1 (21.94)
Teacher 4:	AM	26.8 (5.09)	1.6 (3.17)	24.9 (20.90)
Total:		27.2 (4.88)	1.8 (2.75)	24.2 (19.20)
		School 3		
Teacher 1		23.7 (5.80)	1.5 (3.35)	16.1 (17.85)
Teacher 2		27.3 (9.92)	2.7 (4.18)	34.0 (30.59)
Teacher 3		25.3 (9.54)	2.5 (3.64)	37.8 (34.30)
Total:		25.4 (8.64)	2.2 (3.72)	29.4 (29.71)

Table 3

Total-Group, Between-Class, and Within-Class Correlations of Fall WRAT with Selected Spring Tests

Spring Test	Total-Group	Between-Class	Within-Class		
			Median	Minimum	Maximum
WRAT	.66	.41	.45	.15	.76
Woodcock	.50	.48	.45	.20	.79
Chicago	.61	.15	.40	.22	.77

Table 4

Between-Class Squared Multiple Correlations of Fall WRAT and Classroom Reading Activities with Spring Test Scores

	Spring Score		
Predictors[a]	WRAT	Woodcock	Chicago
WRAT	.165	.231	.022
WRAT & RACT	.821	.407	.641
Unique (RACT)	.656*	.176	.619*
WRAT & D-Time	.767	.507	.600
Unique (D-Time)	.602	.276*	.578*
WRAT & POS-F	.789	.663	.763
Unique (POS-F)	.624*	.432*	.741*

[a]RACT = total number of reading related classroom activities

D-Time = total time spent on decoding instruction

POS-F = total amount of positive feedback

*Significant at .05 level

Table 5

Proportion of Spring Test Score Variance Predictable from Fall WRAT, Classroom, and WRAT by Classroom Interactions

	Spring Test Score		
Predictors	WRAT	Woodcock	Chicago
WRAT	.44	.25	.37
WRAT & Class,	.50	.31	.50
WRAT, Class, & Interaction	.64	.43	.64
Unique (Class)	.06*	.06*	.13*
Unique (Interaction)	.14*	.12*	.14*

*Increase in squared multiple correlation significant at .05 level.

Table 6

<u>Proportion of Spring Score Variance on the Chicago Test Predictable from Fall WRAT, Individual Student Level Reading Variables, and Teacher Effect Codes</u>

Predictors	Squared Multiple R	Change in Squared R
WRAT	.37	.37
WRAT & RACT	.50	.13*
WRAT, RACT, & T	.53	.03*
WRAT & D-Time	.51	.14*
WRAT, D-TIME, & T	.55	.04*

Note. RACT = The total reading related interactions with individual students.

D-Time = The total amount of decoding instructional time received by individual students.

*Increment in squared multiple significant at .05 level.

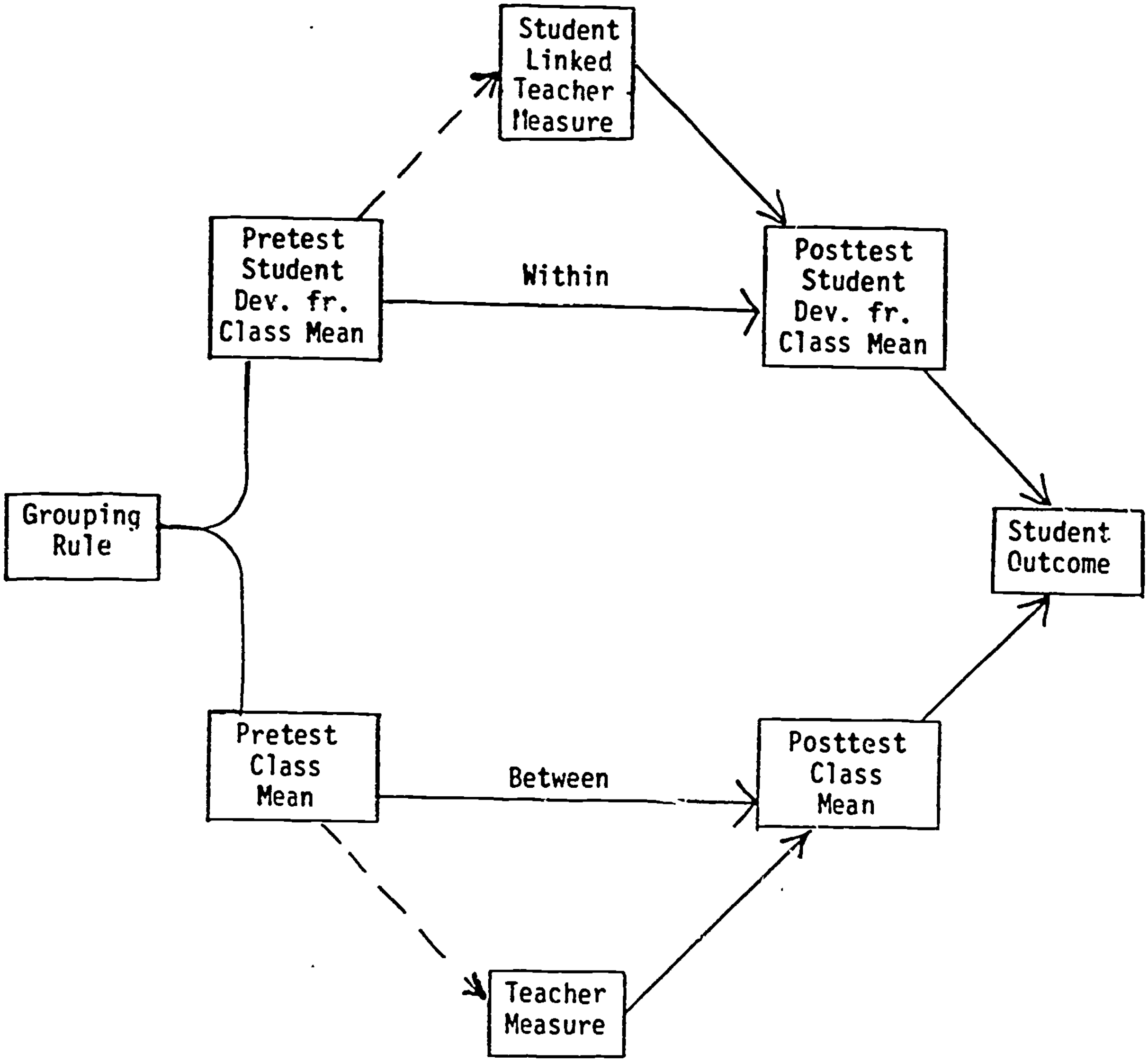

Figure 1

Model of Between and Within-Class Relationships

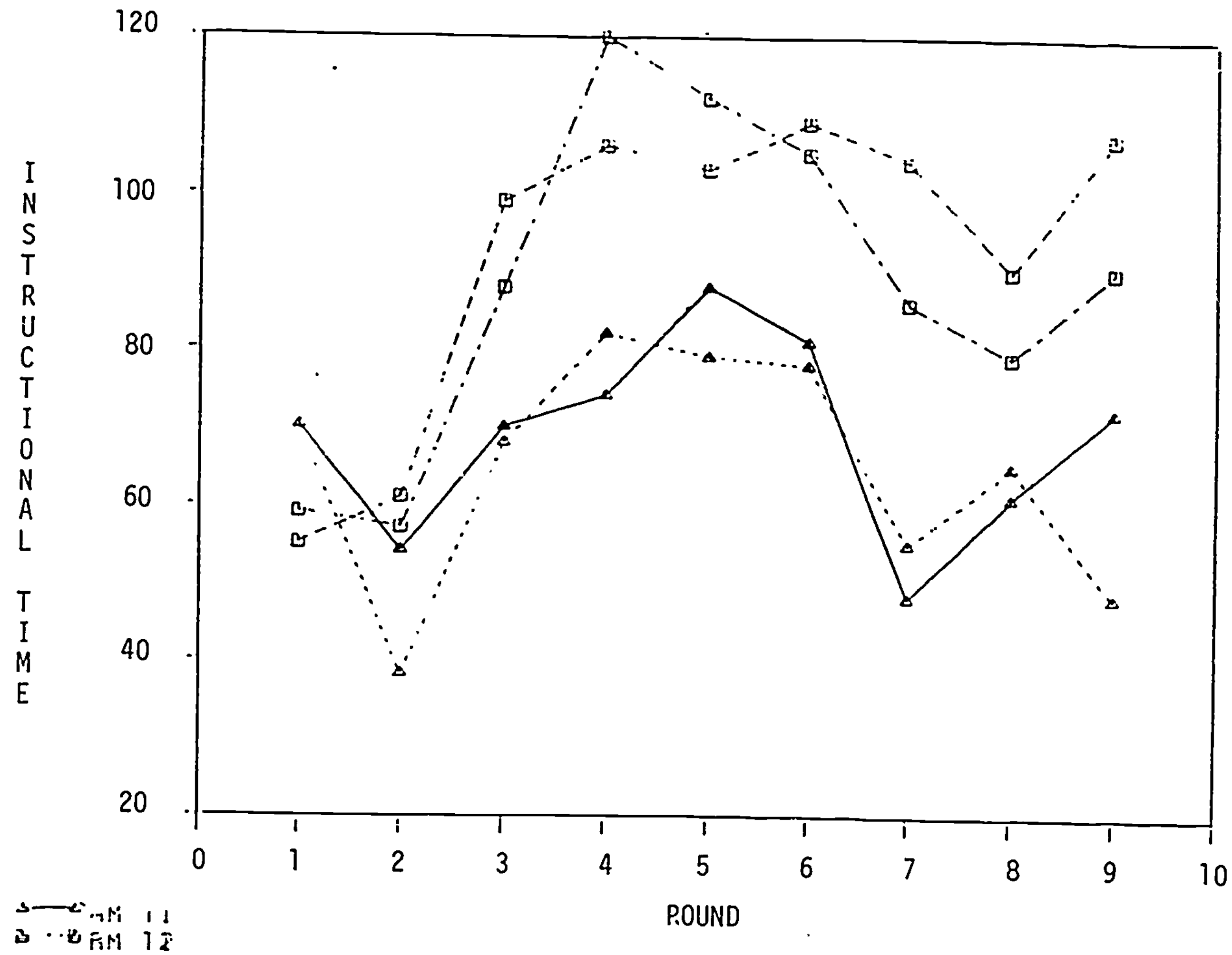

Figure 2

Plots of Instructional Time
by Observational Round for
AM and PM Class of Two Teachers
(□ = school 1, △ = school 2)

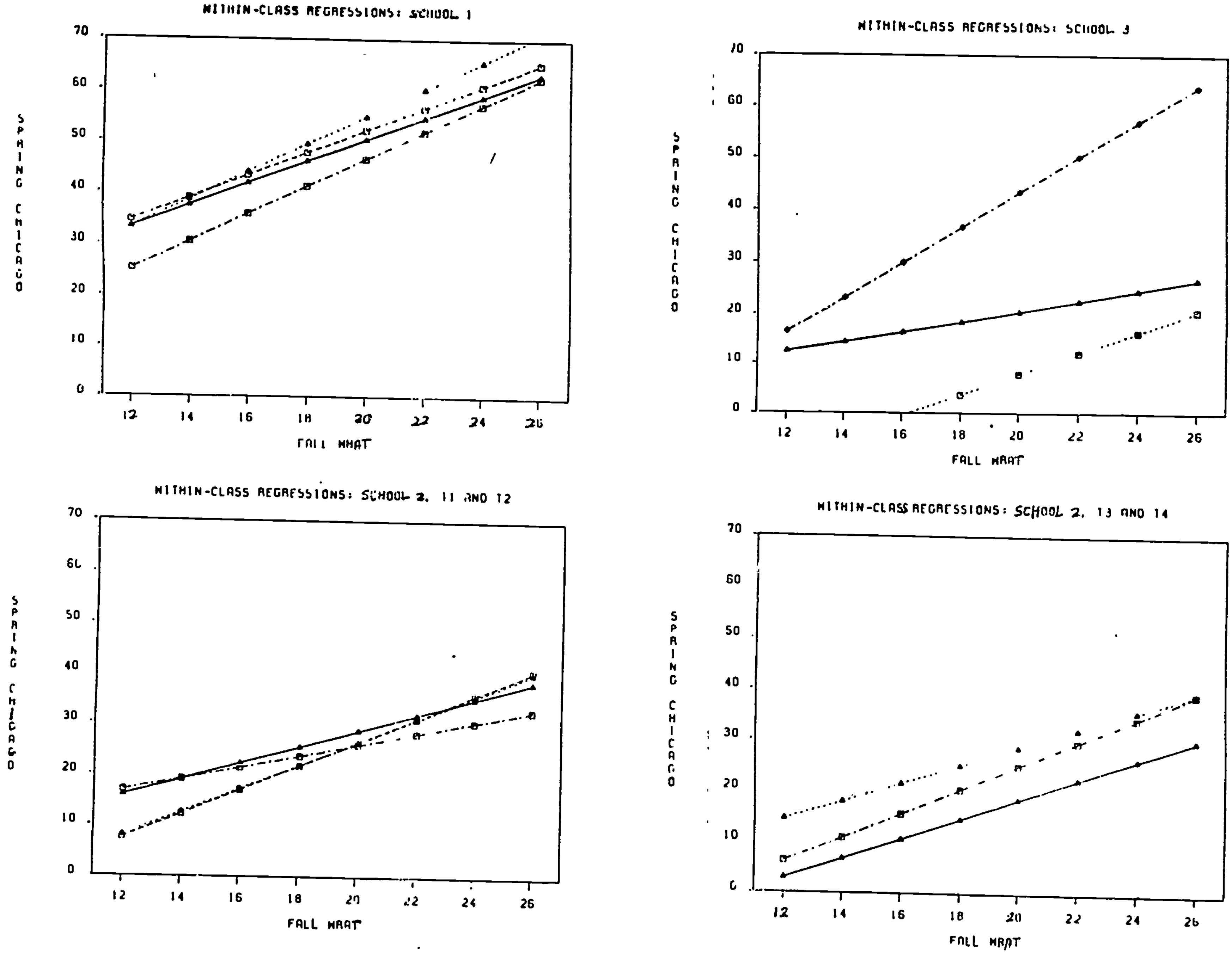

Figure 3
Within-Class Regressions of Spring Chicago on Fall WRAT