DOCUMENT RESUME

ED 125 114 EA 008 442

AUTHOR	Niedermeyer, Fred C.; Fischer, Kathleen B.
TITLE	Developing Exportable Curriculum Supervision for Criterion-Referenced Instructional Programs. Technical Report 49.
INSTITUTION	Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
REPORT NO	SWRL-TR-49
PUB DATE	Feb 74
NOTE	22p.; For a related document, see ED040164 and EA008443
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	Academic Achievement; *Criterion Referenced Tests; Curriculum Development; Educational Assessment; Elementary Education; *Instructional Programs; *Performance Based Education; Principals; Reading Programs; *Supervision; Teachers; Teacher Supervision
IDENTIFIERS	*Instructional Improvement Kits

ABSTRACT

An exportable curriculum supervision system was developed to help supervisors and teachers accomplish three primary functions: to establish class performance expectations, both in terms of pupil achievement and rate of program completion; to monitor actual class performance on a regular, efficient, and systematic basis; and to use suggested strategies to modify instruction and correct observed performance deficiencies. Several versions of the system, termed the Instructional Improvement Kit, were empirically tested to observe the differential effects of program monitoring versus program monitoring plus instructional modification, as well as follow-up teacher meetings. Superior achievement is obtained when a program monitoring system is supplemented by instructional modification strategies that assist supervisors and teachers in analyzing instruction and making changes that have been demonstrated to be generally effective in improving pupil performance. (Author)

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

TECHNICAL REPORT 49

FEBRUARY 1974

DEVELOPING EXPORTABLE CURRICULUM SUPERVISION FOR CRITERION-REFERENCED INSTRUCTIONAL PROGRAMS

Fred C. Niedermeyer and Kathleen B. Fischer

ABSTRACT

An exportable curriculum supervision system was developed to help supervisors and teachers accomplish three primary functions.

1. To establish class performance expectations, both in terms of pupil achievement and rate of program completion.

2. To monitor actual class performance on a regular, efficient, and systematic basis.

3. To use suggested strategies to modify instruction and correct observed performance deficiencies.

Several versions of the system, termed the Instructional Improvement Kit, were empirically tested to observe the differential effects of program monitoring versus program monitoring plus instructional modification, as well as follow-up teacher meetings. Results are presented and discussed.

DEVELOPING EXPORTABLE CURRICULUM SUPERVISION FOR CRITERION-REFERENCED INSTRUCTIONAL PROGRAMS

Fred C. Niedermeyer and Kathleen B. Fischer

The technology for developing instructional programs that can reliably attain important educational outcomes has been described in numerous SWRL publications (Baker & Schutz, 1971). These programs may be rapidly and inexpensively installed in a wide range of school settings by means of exportable teacher training systems (Niedermeyer, 1970). More recent SWRL experience with these programs has indicated that the technology for developing instructional and training systems may also be extended to matters of instructional improvement, conventionally termed "curriculum supervision." The present report describes the development and test of a system, called the Instructional Improvement Kit, that allows supervisors and teachers to identify and correct performance deficiencies and instructional problems as they occur during the year.

THE INSTRUCTIONAL IMPROVEMENT KIT

DESIGN RATIONALE AND CHARACTERISTICS

The Instructional Improvement Kit was designed to help supervisors and teachers accomplish three primary functions:

1. To establish class performance expectations, both in terms of pupil achievement and rate of program completion.

2. To monitor actual class performance on a regular, efficient, and systematic basis.

3. To use suggested strategies to modify instruction and correct observed performance deficiencies.

The importance of these curriculum supervision functions has been evident for some time. It has been observed that teachers and supervisors have difficulty determining satisfactory performance levels and identifying performance problems when they occur. During routine visits by the authors to classrooms using various SWRL programs, it has been noted that some teachers seem satisfied with less-than-minimal class performance, while others impose overly stringent performance standards. Particularly evident has been the teacher variability in scheduling instruction. For various reasons, some teachers simply fail to have all or most of the children complete the recommended number of program units. Consequently, children do not achieve the program outcomes at

a satisfactory level. In most cases supervisors have no means to determine the existence of a problem, much less to initiate or advocate corrective strategies with teachers.

A previous Laboratory study provided some direction for the present development. Implementing rather minimal supervisory type procedures, Sullivan and Niedermeyer (1973) found a significant positive relationship between levels of supervision and end-of-year pupil achievement. The supervisory procedures centered around the notion of class performance monitoring and included: (1) a pacing schedule designed to insure program completion by the end of the school year and (2) means for teachers to periodically report achievement and unit completion data to their supervisors during the year.

The present development effort sought to further operationalize and expand curriculum supervision procedures through the Instructional Improvement Kit. Several features were mandatory in the design of the Kit. First, it was imperative that the materials and procedures incorporated in the system be exportable and usable in a variety of instructional settings. Second, the system should make minimum time demands on both the supervisors and teachers involved. Third, the use of the system was to be self-contained, i.e., supervisors and teachers within an individual school could implement and carry out the system without outside assistance. The system was intended to be a useful prototype for later adaptation within a broader system, e.g., district-wide, computer-managed.

INSTRUCTIONAL VEHICLE

Because of its availability in a large number of schools, the SWRL/Ginn Beginning Reading Program (BRP) was selected as the instructional vehicle for which the prototype Instructional Improvement Kit was developed. The BRP promotes four specific learner outcomes at the kindergarten level: (1) reading words in isolation and in text, (2) saying the sound of various letters or letter combinations (word elements), (3) sounding out and reading new words composed of previously learned word elements (word attack), and (4) reading letter names. The program is divided into ten instructional units of approximately three weeks each. At the end of each unit, the teacher administers a 20-item selected-response test, called a Criterion Exercise (CE). The CE for each unit contains five items on each of the above-listed outcomes. Additional materials are provided to furnish supplementary practice to children whose CE scores indicate less-than-sufficient attainment of unit outcomes.

PROGRAM MONITORING PROCEDURES

The Instructional Improvement Kit developed for use with the BRP was comprised primarily of two components: program monitoring procedures

and instructional modification procedures. These components are described separately.

Program monitoring procedures were designed to assist teachers and supervisors in establishing class performance standards and in periodically monitoring actual performance during the year. Prior to initiation of instruction, teachers and supervisors were asked to determine performance expectations of two types: class achievement and unit completion.

The class achievement standard for the BRP was defined as the proportion of a class scoring 18 or more of the 20 items correct on each unit's Criterion Exercise. This standard was empirically established in an earlier SWRL study (Niedermeyer, 1972). Although teachers and supervisors were free to set their own achievement standards for each unit, SWRL recommendations were given. The SWRL-suggested standard for Unit 1 was 50% of the pupils; for Unit 2, 60% of the pupils; for Unit 3, 70% of the pupils; and for Units 4-10, 80% of the pupils. According to the SWRL-suggested standards, for example, 50% of the pupils in a class would be expected to attain scores of 90% or above on the Unit 1 Criterion Exercise, 60% would be expected to attain this score in Unit 2, 70% in Unit 3, and so on. The gradual gain in achievement levels over the first four units was suggested to allow for increasing familiarity with the program by both pupils and teachers and to avoid unnecessary teacher anxiety over initially low achievement.

Since the intent of the monitoring system is for the supervisor or teacher only to make a decision as to whether or not modification of instructional procedures appears necessary, and since operation of the system should require as little time as possible, it was decided to define the class achievement standard in terms of the total score on each CE rather than in terms of subscores for each of the program outcomes. Once the supervisor or teacher decided instructional improvement strategies should be employed, then scores on each outcome and for each pupil could be examined. Until that time, however, it was felt that the system would be optimally effective if only minimal data were examined for each unit. This is pointed out only because monitoring or "instructional management" systems often present supervisors with an overwhelming amount of data--far more than is required to decide whether or not a teacher requires assistance.

Unit completion standards reflected the pace at which a class progressed through the Beginning Reading Program. To ensure that pupils would complete all ten BRP units and attain the intended program outcomes, two conditions had to be met: (1) program instruction was to be initiated early in the school year, and (2) a regular schedule of unit completion was to be maintained throughout the year. BRP guidelines recommend five weeks for Unit 1 and three weeks for each unit thereafter--a total of 32 weeks. To complete the BRP according to these

guidelines, therefore, a teacher must commence BRP instruction by late September or early October.

At the time achievement standards were to be set, teachers and supervisors were also asked to determine a latest starting date for the program (by which time instruction was to be under way for all children) and to establish projected completion dates for each of the ten BRP units.

Both achievement and unit completion standards were to be recorded by each teacher on a Class Performance Chart provided in the Kit. (Figure 1 provides an example of a partially completed chart. Class performance expectations established by this teacher may be seen at the top of the chart.) During the year, teachers were to complete the chart following each unit's CE; the actual achievement level of t 'ass and the date when the CE was administered were to be recorded. As a class or group completed a CE, the chart was to be submitted to the building-level supervisor for the program (usually the principal). The supervisor could then combine group data and maintain a single Class Performance Chart for the school, listing each class separately. It was suggested that a supervisor allow no more than two units of low achievement or slow pacing to pass without attempting to discuss the problem with the teacher or to assist in any way possible.

Program monitoring procedures (termed "program assessment" in the Kit) were detailed in a 15-minute filmstrip-cassette presentation. After viewing the filmstrip, teachers completed a six-page booklet of practice exercises designed to facilitate skills in completing and interpreting the Class Performance Chart. In the filmstrip, program monitoring procedures were presented primarily as a self-monitoring device for teachers, even though the charts were to be submitted to supervisors (principal) after each unit. The management system was not portrayed as a means for supervisorial checks on instruction, bu in the spirit of teacher-supervisor cooperation and coordination. The clear intent was for teachers to monitor their own progress, requesting or receiving assistance when needed.

INSTRUCTIONAL MODIFICATION PROCEDURES

The second major component of the Kit was instructional modification procedures, designed to help teachers and supervisors analyze and correct performance deficiencies observed through the monitoring system. Instructional modification procedures (termed "instructional improvement" in the Kit) were described in a 20-minute filmstrip-cassette presentation; teacher practice in identifying performance difficulties and in determining appropriate modification strategies was provided in a 13-page booklet of items related to the improvement procedures. Like the program monitoring procedures, the instructional modification procedures were presented primarily as a self-corrective

Latest Starting Date October 9

SWRL BEGINNING READING PROGRAM
CLASS PERFORMANCE CHART

Name Teacher A

Class or Group and Number of Pupils	Unit Number and Number of Weeks Suggested for Instruction									
	Unit 1 (5 wks.)	Unit 2 (3 wks.)	Unit 3 (3 wks.)	Unit 4 (3 wks.)	Unit 5 (3 wks.)	Unit 6 (3 wks.)	Unit 7 (3 wks.)	Unit 8 (3 wks.)	Unit 9 (3 wks.)	Unit 10 (3 wks.)
	Class Achievement Criteria/Expected Completion Date									
	50% / Nov. 3	60% / Nov. 24	70% / Dec. 15	80% / Jan. 24	80% / Feb. 14	80% / Mar 7	80% / Mar 28	80% / Apr 25	80% / May 16	80% / June 6
Group 1 10 pupils	7/10 70% / Oct. 25	8/10 80% / Nov. 10	9/10 90% / Dec. 1	10/10 100% / Dec 20	10/10 100% / Jan. 17					
Group 2 10 pupils	5/10 50% / Nov. 1	6/10 60% / Nov. 22	8/10 80% / Dec. 13	9/10 90% / Jan 19						
Group 3 10 pupils	4/10 40% / Nov. 8	5/10 50% / Dec 1	7/10 70% / Jan 5							
Totals: (achievement only)	16/30 53%	19/30 63%	24/30 80%							

Figure 1. Sample Class Performance Chart

approach for teachers. Cooperation between teachers and supervisors was encouraged, but a careful effort was made to avoid depicting the system as one "imposed" from the administrative level.

The discussion of instructional modification procedures was divided in two parts: one dealing with instructional scheduling (unit completion standards) and one dealing with class achievement. In the filmstrip, procedures for each part were organized around a series of teacher-posed questions. The following questions pertained to problems with instructional scheduling:

- Before children begin a reading program, shouldn't they have more readiness work, such as learning the alphabet and letter sounds, or completing the SWRL Instructional Concepts Program?
- How can I include the Beginning Reading Program in my daily schedule when there are so many other things to do in kindergarten?
- How can I keep up with the suggested three-week-per-unit schedule when there's so much for the children to learn in kindergarten?

Responses to these questions centered around the notions that (1) readiness activities are an integral part of the BRP and that most teachers have found readiness fears unwarranted, (2) reading outcomes deserve high priority in a kindergarten curriculum, and (3) the BRP is designed such that the recommended number of learning activities corresponds to the number of days allowed for completion of each unit.

The following questions were included in the part of the filmstrip pertaining to class achievement deficiencies:

- How can I tell which children are not doing well?
- How can I tell if some children can't succeed in the reading program, even when given additional instruction?
- How can I tell where children are having the most difficulty?
- How can I tell if I am trying to teach too much at one time?
- How do I determine how much practice is necessary for a lesson outcome?

In responding to these questions, the filmstrip directed teachers and supervisors to examine scores on each unit's Criterion Exercise to identify individual children and/or outcomes that seemed to require further attention. Where it was apparent from CE scores that individual children were having difficulty with one or more outcomes, the following

suggestions were given: (1) group these children together and allow them to progress at a slower pace while the rest of the class moves ahead, (2) provide additional practice for these children, by using the SWRL Tutorial or Parent-Assisted Learning Programs, or (3) schedule a staggered day.

On the other hand, where achievement deficiencies were apparent for most or all children on one or more outcomes, it was recommended that teachers: (1) review training materials and procedures with respect to these outcomes, and (2) follow the BRP activity guide, introducing no more than one new lesson outcome per day. A review of lesson procedures was particularly encouraged. One section of the filmstrip extensively reviewed correct procedures for teaching the word attack outcome, as achievement scores had always been lowest in this area, and improper word attack instruction had been frequently observed by the authors in classrooms.

The filmstrip also encouraged teachers to examine a child's unit test scores over two or three units before deciding that he or she wasn't "ready" for reading instruction. Finally, the filmstrip suggested that supervisors or other teachers observe a BRP lesson, using the Lesson Observation Sheet provided in the follow-up practice booklet. This observation sheet related to such general instructional procedures as appropriate practice, rate of response, and number of individual response opportunities. Practice in using the observation sheet was included in the practice booklet.

IMPLEMENTATION OF THE INSTRUCTIONAL IMPROVEMENT KIT

To allow a principal or other school supervisor to implement the Instructional Improvement Kit, a Coordinator's Guide was included with the Kit materials. This guide detailed procedures for (1) scheduling and conducting a one-hour workshop with kindergarten teachers on program assessment (monitoring) procedures and a one-hour workshop on instructional improvement (modification) procedures, and for (2) monitoring Class Performance Charts and assisting teachers during the remainder of the year. The first workshop was to be held prior to mid-October, the second to be held once teachers initiated the BRP but prior to mid-November. The guide sent to each coordinator differed in content according to the school's version of the Kit (see following section).

EMPIRICAL TEST

After development of the Instructional Improvement Kit, several variations were identified within the Kit itself and were empirically tested in a sizable population of kindergarten classes.

TRYOUT PARTICIPANTS

The tryout involved 99 kindergarten classes in 36 elementary schools from eight Southern California districts. Prior to selection for the study, all schools had purchased the SWRL Beginning Reading Program for use during the 1972-73 school year. The total population included over 2,500 kindergarten children, representing a wide range of ethnic groups and socio-economic levels.

VARIATIONS TESTED

The two components examined during the tryout involved the Instructional Improvement Kit itself and Follow-up teacher meetings. The variations tested for each of these components are described below.

Instructional Improvement Kit

Two versions of the Instructional Improvement Kit were tested. The first version consisted of monitoring procedures (PM), i.e., those procedures and materials that pertained to establishing performance standards and to monitoring actual performance during the year. This level of the treatment involved the first filmstrip-cassette and kit materials described earlier (see the Program Monitoring Procedures section).

The second version of the Instructional Improvement Kit involved both program monitoring and instructional modification (PM & IM). In addition to the PM procedures and materials, this level of the treatment included means to analyze and correct performance difficulties (see the Instructional Modification Procedures section). Materials for this level consisted of the complete Instructional Improvement Kit, i.e., two filmstrip-cassette presentations and corresponding materials.

The two versions of the Kit were designed to isolate the effects of monitoring (PM) versus monitoring and instructional modification (PM & IM). It was believed that the more comprehensive treatment (PM & IM) would prove more effective in terms of class performance in the BRP because teachers and supervisors would have means not only to identify performance deficiencies, but to correct them as well.

Follow-up Meetings

The second component tested, Follow-up Meetings, was intended to facilitate the implementation of the Instructional Improvement Kit's procedures. By structuring an environment in which teachers and

supervisors could gather periodically to discuss BRP performance, it was felt that modification strategies were more likely to be formulated and initiated. Without this structure, it was uncertain whether other forms of teacher interaction or teacher-supervisor cooperation would occur.

Accordingly, with Follow-up Meetings (FM), school coordinators or supervisors (usually the principals) were asked to schedule and conduct three Follow-up Meetings during the year, after Units 1, 3, and 6 of the BRP. (These three meetings were in addition to the one or two meetings held at the beginning of the school year to introduce and work through the Instructional Improvement Kit.) The teacher meetings were to serve as a forum for discussion of difficulties with the program. Current BRP performance was to be reviewed, and modification strategies were to be reviewed. Each FM supervisor was asked to submit a report of the meeting to the experimenters. Summary forms and stamped, SWRL-addressed envelopes were provided for this purpose. Supervisors and teachers not under the FM condition were not asked to conduct or participate in Follow-up Meetings.

TRYOUT DESIGN

The previously described variations were combined to form four tryout groups: Program Monitoring Only (PM), Program Monitoring Only combined with Follow-up Meetings (PM; FM), Program Monitoring plus Instructional Modification (PM+IM), and Program Monitoring plus Instructional Modification combined with Follow-up Meetings (PM+IM; FM).

Twenty four of the 36 schools were then randomly assigned to one of the four groups. The remaining 12 schools were designated as BRP-only comparison schools.

DATA SOURCES

Pupil Achievement Posttest

A 40-item, constructed-response posttest was used to assess end-of-year achievement on the outcomes of the BRP. The test consisted of 10 items on each of the four program outcomes. It was individually administered and required the child to make oral responses to printed words or letters. The test had been developed prior to the study. Reliability coefficients for BRP tests of this same format had consistently been found to be .90 or higher. (See, for example, Niedermeyer, 1969.)

Teacher Questionnaire

At the time of posttesting, teachers were asked to complete a short, two-page information sheet, or questionnaire. This questionnaire provided information as to the number of BRP units completed by each child tested, as well as to the extent to which the appropriate treatment condition had been implemented, e.g., "Did you view the filmstrip, 'Program Assessment Procedures' (explained how to fill in the Class Performance Chart after each unit)?"

Supervisor Questionnaire

Supervisors (principals) were sent short reaction sheets similar to the teacher questionnaires. The supervisors' questionnaire assessed the extent to which respective treatment conditions had been carried out, e.g., "How often did you see the updated Class Performance Chart for each class during the year?"

Completed Class Performance Charts

In addition to returning completed reaction sheets at the end of the study, principals were asked to submit any completed Class Performance Charts that they and their teachers had maintained during the year.

Meeting Summaries

Supervisors in the 12 schools under the Teacher Meeting condition were provided with forms on which to summarize meetings held during the year with teachers. These were mailed to the Laboratory after each meeting.

Comparison School Questionnaires

Although teachers and supervisors in the 12 BRP-only comparison schools did not receive instruction improvement or follow-up meeting materials, they were asked to complete questionnaires indicating the nature of any management tasks that might have been performed independently.

PROCEDURES

Schools were selected and assigned to tryout groups in late August and early September of 1972. A check on assignment bias among schools was made on the basis of a measure of prior status, the schools' first-

grade median score on the 1971-72 state reading achievement test. An analyses of variance was performed, using the class as the unit of analysis. No significant differences between the tryout groups were revealed.

No contact was made with the 12 comparison schools until the end of the school year. In early September, the principals of each of the 24 tryout schools were contacted by phone and invited to participate in the field test. All principals indicated their willingness to cooperate and agreed to carry out the tryout as specified. Each principal was encouraged to serve as supervisor for the tryout in his or her school, although it was made clear that this responsibility could be delegated to another person on the staff, e.g., a teacher or vice principal. No further contact was made with the schools until the end of the school year. The 12 teacher meeting schools, however, were to submit meeting summaries by mail

During the year, three of the 24 tryout schools sent letters requesting to be dropped from the study due to lack of time. These letters were received between November and February. One school dropped from each of the following groups: (PM), (PM; FM), and (PM+IM).

In late April, all 33 schools (12 comparison and the 21 remaining tryout schools) were contacted to schedule posttesting. Posttesting was conducted in mid-May of 1973. In each class, a trained SWRL staff member individually administered the 40-item posttest to a random sample of 10 children, and the teacher completed the questionnaire. Letters containing the supervisor questionnaire and requesting completed Class Performance Charts were sent to supervisors in late May. Posttest scores were obtained from all classes, and questionnaires were obtained from all but four of the 33 principals and three of the 89 teachers.

RESULTS

PUPIL POSTTEST PERFORMANCE

Since implementation of the BRP and use of the materials and procedures in the Instructional Improvement Kit were conducted by the classroom teacher, class means rather than pupil means formed the unit of analysis throughout the study. Mean scores on the 40-item reading posttest for the 60 classes comprising the four tryout groups are shown in Table 1. Scores for the 28 BRP-only comparison classes are shown in Table 2. These tables reveal that instructional modification combined with program monitoring seemed to have the most impact on pupil achievement. The mean scores were 27.9 and 27.6 for the two PM+IM groups and 19.3 and 25.0 for the two PM groups. The mean comparisons between the two follow-up meeting groups and the two groups with no follow-up meetings are in the expected direction, but are not

Table 1

Mean Posttest Scores and Number of Units Completed--All Tryout Classes

Group	No. of Classes (& Schools)	Mean Raw Posttest Score (40 possible)	Standard Deviation	No. Units Completed (out of 10)	Standard Deviation
Program Monitoring Only	12(5)	19.3	9.3	7.4	2.7
Program Monitoring; with Follow-up Meetings	14(5)	25.0	5.3	8.4	0.6
Program Monitoring plus Instructional Modification	15(5)	27.9	6.8	7.8	1.5
Program Monitoring plus Instructional Modification; with Follow-up Meetings	19(6)	27.6	6.0	8.0	1.9
Total	60(21)	25.2	---	7.9	---

Table 2

Class Posttest Mean Scores and Number of Units Completed--All BRP Only Comparison Classes

Group	No. of Classes (& Schools)	Mean Raw Posttest Score (40 possible)	Standard Deviation	No. Units Completed (out of 10)	Standard Deviation
Non-BRP Comparison Classes	28(12)	20.2	10.2	5.9	2.8

as substantial in magnitude as that observed for PM+IM. All tryout groups except the PM Only group scored substantially higher on the posttest than the comparison group (Table 2).

Class means on the posttest for the four tryout groups in Table 1 were subjected to a two-way analysis of variance. The results of this analysis appear in Table 3. As may be seen, the Instructional Improvement Kit but not the follow-up meetings produced a statistically significant effect.

Table 3

Two-Way Analysis of Variance of Class Posttest Means--All Tryout Classes

Source	SS	df	MS	F
Instructional Improvement Kit (A)	456	1	456	9.1*
Follow-up Meetings (B)	103	1	103	2.1
A x B	1[illegible]	1	130	2.6
S/AB	2806	56	50	

*p < .01

UNITS COMPLETED

Tables 1 and 2 also present the mean number of BRP units completed (out of 10) for each of the four tryout groups and the comparison group. Very little variability can be detected among the four tryout groups in Table 1. (An ANOVA revealed no significant main effects or interaction.) The mean number of units completed by all groups was 7.9 or nearly 80% of the program. The BRP-only comparison classes, however, completed an average of only 5.9 units during the year.

VARIABILITY IN USE OF THE KIT AND IN FOLLOW-UP MEETINGS

One of the problems with assessing the effects of programs or innovations in school settings is that it is often left unclear to what extent the programs are "real" or "fictional" (Charters & Jones, 1973). The Instructional Improvement Kit described in this study, for example, was "real" in classes where (1) teachers and principals actually met and worked through the kit materials, (2) the teachers

set performance and pacing standards and completed the Class Performance Chart during the year, (3) principals examined and updated charts on a regular basis, and (4) principals and teachers formulated and tested instructional modifications when problems became evident. The kit is "fictional" or only partly real in classes where any or all of these things did not occur. Too often group data representing the performance of a particular program or innovation describe both real and fictional situations so as to lower estimates of what the real program is capable of producing. Thus, it is informative and useful to assess the extent to which thorough program implementation occurred and to examine the performance data correspondingly.

Several data sources were used to assess the extent to which the Instructional Improvement Kit and the follow-up meetings were actually implemented in the 60 tryout classes in 21 schools. These sources included teacher and principal questionnaires, completed Class Performance Charts (Fig. 1), and reports of follow-up meetings.

An examination of these data sources revealed that classes in seven of the 21 schools did not carry out their respective curriculum supervision procedures thoroughly enough to be considered implemented. A listing of all 60 tryout classes, indicating which implemented the program throughout and which did not, appears in Table 4. For each school considered not implemented, reasons for arriving at this judgment are indicated. (Table 4 also contains the mean posttest score and number of units completed for each of the 21 treatment schools and 12 BRP-only comparison schools.) Generally, the most frequent indicators of lack of implementation were the failure of the principal to monitor class performance or hold follow-up meetings during the year, following the initial training meetings.

Table 5 shows the posttest and unit completion class means for classes in the schools that were found to have thoroughly implemented their respective curriculum supervision procedures. When compared to Table 1, it may be seen that Table 5 shows generally higher performance but has the same patterns of groups differences. PM+IM groups still show the highest achievement. An ANOVA of the data in Table 5 revealed the identical pattern of statistical differences as was found for the complete data in Table 1.

ATTITUDES TOWARD THE KIT

As previously mentioned, three of the original 24 tryout schools selected for the field test requested to be dropped; all gave lack of time as a reason. Of the remaining 21 schools, only one indicated displeasure with the curriculum supervision procedures. Again, the reason was that they required too much time. No other negative comments were

Table 4

School Posttest Mean Scores, Number of Units Completed, and Extent of Implementation by Each Treatment and Comparison Group

Group	School No.	No. of Classes	Mean Posttest Score	No. Units Completed	Extent of Implementation
Program Assessment; No Teacher Meetings	1	3	9.4	6.9	Implemented (low CE scores on CPC*)
	2	2	22.5	7.7	Not implemented (CPC's not returned. No monitoring)
	3	2	9.7	2.3	Not implemented (Teacher's didn't see filmstrip. No CPC's returned)
	4	2	33.2	9.5	Implemented
	5	3	26.6	9.8	Implemented
Program Assessment; Teacher Meetings	6	2	30.0	8.4	Implemented
	7	3	21.7	7.2	Implemented
	8	2	20.3	8.0	Implemented
	9	3	25.8	8.6	Implemented
	10	4	25.9	9.0	Not implemented (No teacher meetings; no monitoring)
Program Assessment + Instructional Improvement; No Teacher Meetings	11	4	34.2	8.1	Implemented
	12	3	19.1	6.4	Implemented
	13	3	30.3	9.0	Implemented
	14	2	19.1	5.5	Not implemented (II filmstrip not seen. No CPC's received)
	15	3	33.1	9.0	Implemented

*CPC = Class Performance Chart

Table 4--continued

Group	School No.	No. of Classes	Mean Posttest Score	No. Units Completed	Extent of Implementation
Program Assessment + Instructional Improvement; Teacher Meetings	16	3	25.5	7.0	Implemented
	17	2	13.6	3.6	Not implemented (No CPC's received. No meetings or monitoring)
	18	6	33.3	10.0	Implemented
	19	3	25.8	7.4	Implemented
	20	2	26.5	8.0	Not implemented (No teacher meetings or monitoring)
	21	3	29.4	9.4	Not implemented (No teacher meetings or monitoring)
Comparison Schools	22	2	13.8	4.7	
	23	2	25.4	7.0	
	24	2	30.9	8.0	
	25	3	12.7	3.2	
	26	1	23.5	5.9	
	27	3	24.5	6.5	
	28	3	9.5	1.5	
	29	3	23.8	6.1	
	30	2	36.4	10.0	
	31	2	6.5	3.7	
	32	3	32.9	9.7	
	33	2	9.8	5.9	

Table 5

Mean Posttest Scores and Number of Units Completed--Only Those Tryout Classes Fully Implementing the Instructional Improvement Kit

Group	No. of Classes (& Schools)	Mean Raw Score (40 possible)	Standard Deviation	No. Units Completed (out of 10)	Standard Deviation
Program Monitoring Only	8(3)	21.0	9.9	8.7	1.6
Program Monitoring; with Follow-up Meetings	10(4)	24.6	4.4	8.2	0.5
Program Monitoring plus Instructional Modification	12(4)	30.2	5.8	8.3	1.1
Program Monitoring plus Instructional Modification; with Teacher Meetings	12(3)	29.7	3.7	8.4	1.4
Total	42(14)	27.0	---	8.4	---

received from any data sources with regard to the Instructional Improvement Kit. A number of teachers did indicate that they felt the SWRL-suggested performance criteria were too high, but since teachers were able to establish their own criteria, this did not seem to be a serious problem.

Several teachers and principals volunteered positive comments about the curriculum supervision procedures on their questionnaires. These comments generally were related to the quality of the materials (filmstrips, charts) and to the value of the monitoring and modification procedures themselves.

DISCUSSION

The results of this study suggest that an exportable curriculum supervision system, such as the BRP Instructional Improvement Kit, can significantly improve pupil performance in a criterion-referenced instructional program. In all but the lowest-level tryout conditions

(Program Monitoring Only), mean posttest achievement scores were substantially higher than those of the program-only comparison group. In addition, tryout groups completed about two more program units than did the comparison group.

The monitoring system itself may serve as one explanation for differences. Program monitoring procedures presented in the Kit furnished teachers with procedures and suggestions for setting reasonable performance expectations for their kindergarten pupils. By working toward these expectations, teachers could conduct on-going self-evaluation of their efforts during the school year. In this way, teachers would know when to focus their attention on improving instruction, and could be more conscious of existing problems than they might have been without any guidelines for evaluation.

At the same time, the monitoring system seemingly drew supervisory personnel into a more active dialogue with teachers on program progress. Principals and coordinators were kept informed of current performance levels in the program and could theoretically direct their efforts toward promoting improvement in difficult areas.

Significantly greater gains in pupil performance were observed with the addition of instructional modification or improvement strategies. Again, the mere focusing of attention onto problem areas may have sufficiently raised the consciousness of teachers and supervisors to avoid possible pitfalls in program implementation. On the other hand, the initiation of improvement strategies themselves may have accounted for the gains made by the PM+IM groups. Of the strategies presented in the improvement filmstrip, it does not appear that those relating to unit completion schedules had any measurable effects, since mean numbers of units completed for the PM and PM+IM groups do not differ significantly. It is more likely that procedures related to analyses of criterion exercise scores and strategies to correct achievement deficiencies had greater impact. In particular, procedures related to word-attack instruction may have been of considerable consequence in improved pupil achievement scores.

When considering the effectiveness of an instructional improvement system of the type developed here, it should be noted how important it is for the supervisor (in this case each school principal) to thoroughly implement, monitor, and participate in the system. When the 60 tryout classes were examined to determine the extent to which they implemented the Instructional Improvement Kit, it was found that all of the classes in seven of the 21 schools were essentially "untreated" because the supervisor failed to carry out the Kit's procedures with the teachers. It may be that more consistent school-by-school implementation of an instructional improvement system would be obtained if an administrator at the district level implemented the system through the principals and asked each principal to regularly submit brief school progress reports. Principals in this tryout were presented with no general

rationale or had little extrinsic incentive, other than pleasing SWRL, to implement the Instructional Improvement Kit. Hopefully, they had considerable intrinsic motivation for curriculum supervision tasks. However, when pressed for time, meaningful extrinsic motivation, such as reporting to a district-level administrator, may be effective. Minimally, the rationale for and expected desirable consequences of their participation should be carefully explained.

It is also possible that in many cases leadership for the curriculum supervision should not be restricted to the principal. Curriculum consultants, district coordinators, or teachers themselves are other supervisor options schools should consider when implementing this type of system.

It should also be noted that implementation of curriculum supervision procedures for a kindergarten reading program is not as likely to generate as much administrator interest and effort as more traditional areas. It may be that principals would be much more willing to be instructional leaders in a reading program for grades one through six, where much more has been expected in terms of teacher effort and pupil achievement.

In short, it appears possible to develop exportable curriculum supervision that can positively affect the rate of program completion and pupil achievement. Superior achievement is obtained, however, when a program monitoring system is supplemented by instructional modification strategies that assist supervisors and teachers in analyzing instruction and making changes which have been demonstrated to be generally effective in improving pupil performance.

REFERENCES

Baker, R. L., & Schutz, R. E., eds. Instructional product development. New York: Van Nostrand Reinhold, 1972.

Charters, W. W., Jr., & Jones, J. E. On the risk of appraising non-events in program evaluation. Educational Researcher, 1973, 2 (11), 5-7.

Niedermeyer, F. C. Parent-assisted learning. Technical Report 19, 1969, SWRL Educational Research and Development, Los Alamitos, California.

Niedermeyer, F. C. Developing exportable teacher training for criterion-referenced instructional programs. Technical Report 22, 1970, SWRL Educational Research and Development, Los Alamitos, California.

Niedermeyer, F. C. Differential effects of individual and group testing strategies in an objectives-based instructional system. Journal of Educational Measurement, 1972, 9, 199-204.

Sullivan, H. J., & Niedermeyer, F. C. Pupil achievement under varying levels of teacher accountability. Professional Paper 28, 1973, SWRL Educational Research and Development, Los Alamitos, California.