DOCUMENT RESUME

ED 102 481 95 CG 009 582

AUTHOR Klein, Roger D.; Schuler, Charles F.
TITLE Increasing Academic Performance through the Contingent Use of Self-Evaluation.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Apr 74
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; Elementary Education; *Individualized Instruction; *Learning; *Reinforcement; Research Projects; *Self Evaluation; Speeches
IDENTIFIERS *Individually Prescribed Instruction

ABSTRACT

The present study was designed to determine if students would improve their test performance in order to earn the right to self-evaluate their daily tasks, and then, whether they would maintain high test performance in order to keep this privilege. A second purpose was to demonstrate a practical way in which teachers could effectively reduce the amount of time they spent evaluating students' daily tasks, without adversely affecting student test performance. Subjects were students in two third-grade classrooms in an inner-city school. Classrooms made use of the Individually Prescribed Instruction Math Program designed by the Learning Research and Development Center, wherein students, to master a skill, must complete workbook pages and pass a test which measures skill performance. The variable manipulated was the opportunity to self-evaluate workbook performance. Self-evaluation was defined as a set of behaviors which led to a decision by a child to take a test. Children were told that if they passed the first test in their present skill they would be permitted to evaluate all workbook pages for their next skill. A child who attempted a second test and failed lost his self-evaluation privilege but could re-earn it by passing the first test in the next skill. Results of the study indicate that the introduction of the contingent self-evaluation procedure produced improved academic performance for the majority of children in the two classrooms. More importantly, it was demonstrated that children could perform at a high level while sharing responsibility for managing their own behavior. (Author/CKJ)

INCREASING ACADEMIC PERFORMANCE THROUGH THE CONTINGENT USE OF SELF-EVALUATION

Roger D. Klein and Charles F. Schuler

Learning Research and Development Center
University of Pittsburgh

Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April 1974.

The research reported herein was supported by the Learning Research and Development Center, supported in part by funds from the National Institute of Education (NIE), United States Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of NIE and no official endorsement should be inferred. We wish to thank third-grade teachers Candace Ganea and Beverly Lucas as well as educational specialist, Jessie Mike of Manchester Elementary School in Pittsburgh, Pennsylvania, for their full cooperation during the course of this investigation. We also wish to thank Helen D. Klein and Judy Light for their very valuable suggestions on the design of the investigation.

Abstract

The opportunity to self-evaluate daily work performance was made contingent upon test-passing behavior in two individualized third-grade classrooms. A multiple baseline design demonstrated that when the contingent self-evaluation phase was introduced, test performance improved markedly for a majority of the children over levels observed during a teacher evaluation condition. Students showed a higher percent of tests passed, increased mean test scores, and a markedly lower mean rate of daily failures and tests attempted. The mean daily rate of tests passed decreased minimally. It was suggested that students could both improve their academic performance and manage their own behavior without the use of external reinforcers or extensive teacher supervision.

INCREASING ACADEMIC PERFORMANCE THROUGH THE CONTINGENT USE OF SELF-EVALUATION

Roger D. Klein and Charles F. Schuler

Learning Research and Development Center
University of Pittsburgh

Increasing interest has been shown in the use of self-control techniques to modify academic performance in the classroom. Lovitt and Curtiss (1969) found that a higher level of academic performance was produced when contingency requirements were student-determined rather than teacher-determined, and Glynn (1970) reported that student-specified reinforcement was as effective as experimenter-imposed reinforcement in accelerating academic behavior.

More recently, Knapczyk and Livingston (1973) reported that students improved their reading performance, relative to a baseline phase, when the students' ratings of reading accuracy were exchanged for tokens and backup reinforcers. Similarly, Glynn, Thomas, and Shee (1973) showed that students' on-task levels remained as high during a self-recording and self-reinforcing phase as they had when recording and reinforcement were externally controlled. Both Knapczyk and Livingston, and Glynn et al., also concluded that the students were reasonably accurate in their self-recording.

Each of these studies was characterized by the use of relatively powerful incentive systems in which the students were permitted to self-determine and/or self-administer reinforcement. While it certainly is of interest to examine the extent to which students can manage external reinforcement systems, it would also appear desirable to develop self-control procedures, for classroom use, which would not depend regularly upon backup reinforcers to insure their success.

One alternative approach has been simply to train students to self-record a specific behavior, and then evaluate the effectiveness of self-recording in altering the behavior. Broden, Hall, and Mitts (1971) reported the acceleration of study behavior through the use of student self-recording. A second approach has been to require students to evaluate their own behavior, relative to a given standard, at the end of a specified time period. An assesement is then made of the effects of the self-evaluation procedure upon the student's behavior. No self-evaluation study has yet been conducted in which the target behavior has been academic performance. However, Santogrossi, O'Leary, Romanczyk, and Kaufman (1973) had extremely disruptive hospitalized adolescents self-evaluate their inappropriate behavior every 15 minutes with the use of a rating scale. They concluded that self-evaluation alone did not reduce disruptiveness.

A third approach might be to use self-recording, self-evaluation, or some related procedure in a contingent fashion. It has been suggested by several investigators that self-recording/evaluation might function as a reinforcing activity (Kaufman & O'Leary, 1972; Knapczyk & Livingston, 1973). Potentially, a classroom situation could be designed in which students would earn the opportunity to self-record/evaluate, contingent upon their performance of a desired behavior.

The present study was designed to determine if students would improve their test performance in order to earn the right to self-evaluate their daily tasks, and then, whether they would maintain high test performance in order to keep this privilege. A second purpose of the study was to demonstrate a practical way in which teachers could effectively reduce the amount of time they spent evaluating students' daily tasks, without adversely affecting student test performance.

Method

Subjects and Setting

The subjects were the students in two third-grade classrooms in an inner-city public school in Pittsburgh, Pennsylvania. One classroom (A) contained 26 students and the other (B), contained 24 students. The experiment was conducted in each classroom during the math period, which lasted from 10:00 until 10:45 a.m., five days per week. There were two adults in the classroom, a teacher, and an assistant teacher.

Daily Program

The classrooms made use of the Individually Prescribed Instruction (IPI) Math Program designed by the Learning Research and Development Center. IPI Math is broken into ten content areas (e.g., Addition Subtraction, Division) and each content area is broken down into levels of difficulty (e.g., A, B, C). The intersection of a content area and a level of difficulty (e.g., Division Level B) constitute a unit of work. In turn, each unit is divided into a number of skills, and when a child enters a unit he is pretested to determine which skills he needs to learn.

To master a skill, the student must complete workbook pages and pass a test which measures his skill performance. A score of 85 percent or above, is considered passing. If a child passes a test, he goes on to workbook pages in a new skill. If a child fails a test, he is assigned new workbook pages and a second, different test. Under ideal conditions a child would pass the first test in each skill on a regular basis. The present study was concerned only with the first test in each skill, and specifically, with increasing the percentage of these tests that were passed.

Independent Variable

The variable manipulated in the study was the opportunity to self-evaluate workbook performance. Self-evaluation was defined as a

set of behaviors which led to a decision by a child to take a test. The set of behaviors which produced this decision varied from child to child. The behaviors included the choice to use or not to use scoring keys to evaluate workbook pages, and the choice of how to use the scoring keys (e.g., complete work and then check it using the keys, correct errors after examining the keys, copy the answers on to the workbook pages without actually attempting the items). For any given child the behaviors used to self-evaluate performance may have varied from day to day. No attempt was made to determine the method of self-evaluation used by any child.

Dependent Variable

The dependent variable was performance on the first test assigned after completion of a set of workbook pages (a skill). Five measures of test performance were examined. They were: (1) percent of test items correct, (2) percent of tests passed, (3) number of tests passed per day, (4) number of tests failed per day, and (5) number of tests taken per day.

Response Scoring and Reliability

All tests were scored by the assistant teacher with a set of scoring keys. During each week of the experiment the second author randomly selected eight to ten tests in each classroom and, using scoring keys, checked the scoring done by the assistant teacher. Reliability in each case was 100 percent.

Procedure

A multiple baseline design was used in which baseline data were collected for two weeks in classroom A and for three weeks in classroom B. The experimental treatment lasted for seven weeks in classroom A and six weeks in classroom B.

Teacher evaluation. The regular classroom procedure was followed in which the teacher circulated among the children asking and answering questions, and evaluating workbook pages. Children were required to score 100 percent on each day's workbook pages before proceeding to either new workbook pages or a test. All tests were completed by children at their desks. On the last day of this phase, the students were trained to self-evaluate their workbook performance; they were shown how to compare several mock workbook pages with a set of scoring keys. It was suggested to the children that they circle their incorrect answers, when using the keys, in much the same way the teacher did when she had evaluated their pages.

Contingent self-evaluation. The children were told that if they passed the first test in their present skill they would be permitted to evaluate all workbook pages for their next skill. Scoring keys were duplicated so that each child could evaluate his own work at his desk. The teacher suggested to the children that if they had earned the self-evaluation privilege, they should review their answers with the scoring keys, correct their errors, and maintain the 100 percent accuracy criterion which operated during teacher evaluation. It was also suggested to the children that they attempt each arithmetic item and refrain from copying the scoring key directly on to their workbook pages. There was, however, no attempt to enforce any of these suggestions. As a result, children who had earned the self-evaluation privilege were free to self-evaluate in a variety of ways, proceed at their own pace, and attempt to master their next skill test at any point in time.

If a child who was self-evaluating attempted his next test and failed, he lost his self-evaluation privilege. He could re-earn the privilege by passing the first test in the next skill. Until that point, however, the teacher evaluated his daily assignment.

Results

The results are presented only for those students in classrooms A and B who attempted at least three tests during each phase of the study. There were 13 students in classroom A and 12 students in classroom B who met this requirement.[1]

Figure 1 presents weekly data on the percent of tests passed and the mean percent of items correct for both classrooms during the two phases of the study. In classroom A the mean percent of tests passed during the two-week teacher evaluation phase was 65.2 percent, and the mean percent of items correct was 86.6 percent. These averages increased, respectively, during contingent self-evaluation, to 88.6 percent and 93.8 percent. In classroom B the mean percent of tests passed during the teacher evaluation phase was 53.7 percent, while the mean percent correct on tests was 74.7 percent. When contingent self-evaluation was introduced the means increased to 74.6 percent for percent of tests passed, and 87.6 percent for percent of items correct. It should be noted that the two-week decline during weeks six and seven for classroom

---

[1] It was decided that three tests per child during each phase would be the minimum number of tests needed to provide a reliable measure of a student's test performance. However, an analysis of the data which included all students attempting at least one test per phase produced almost identical results. In classroom A, for the 19 students attempting one or more tests, the percent of tests passed during teacher evaluation was 64.9 percent, and the mean percent of items correct was 85.0 percent. During contingent self-evaluation these figures increased to 88.6 percent and 93.8 percent, respectively. In classroom B, where 22 students attempted one or more tests, the percent of tests passed and the mean percent of items correct during teacher evaluation were 58.8 percent and 75.6 percent. When contingent self-evaluation was introduced the respective means increased to 74.0 percent and 86.4 percent. All students in the two classrooms who attempted one or more tests in each phase earned the opportunity to self-evaluate at some point during the treatment condition.

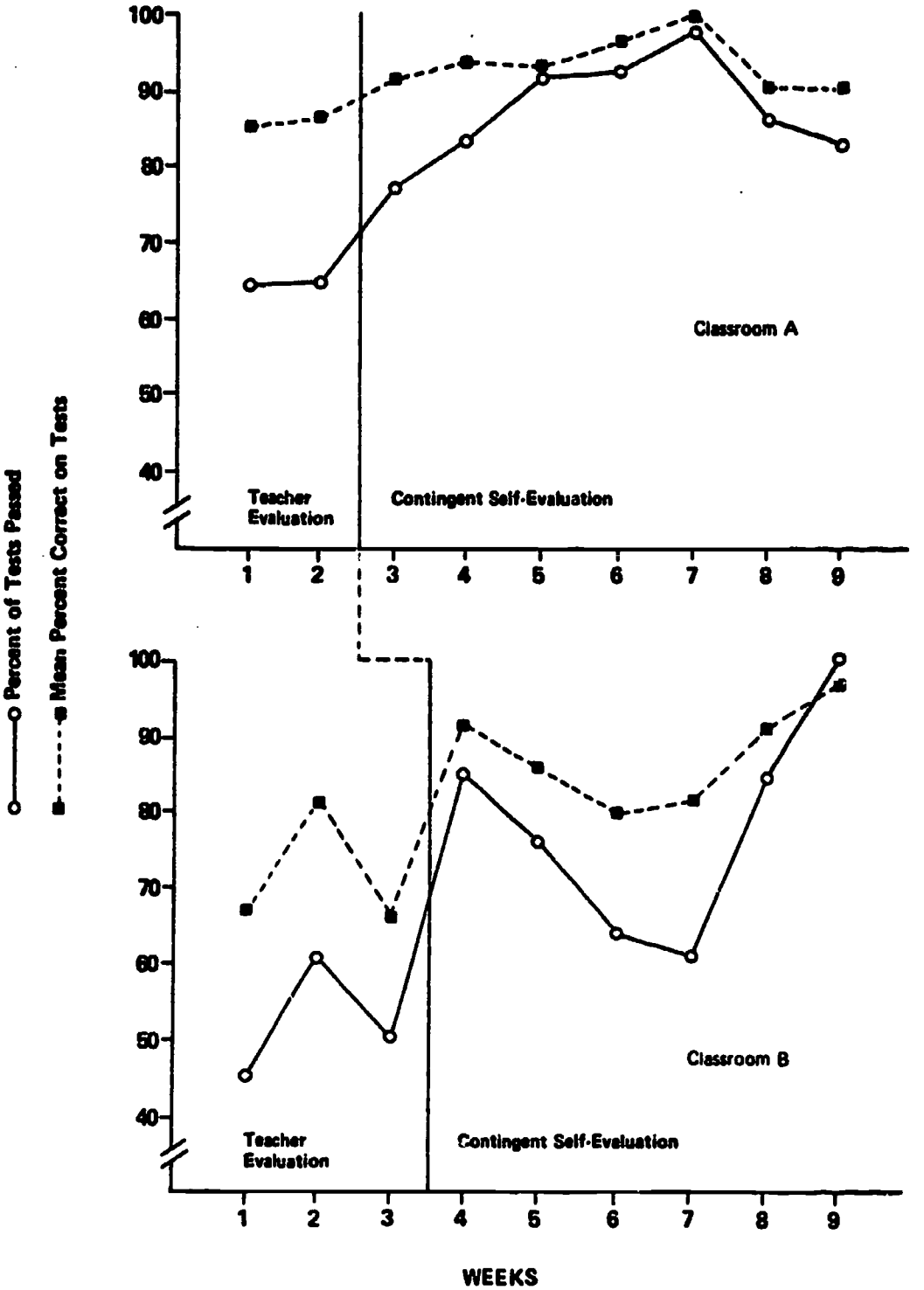

Figure 1. Percent of tests passed and mean percent correct on tests per week for classrooms A and B.

B was not a group phenomenon. During that time two students showed a large increase in the number of tests they attempted. Over 80 percent of their attempts, however, resulted in failure, thus suppressing the group average for these two weeks.

Table 1 presents mean data for both classrooms on tests passed per day, tests failed per day, and tests taken per day during each phase of the investigation. In each classroom there was a slight decrease in the mean number of tests passed per day when contingent self-evaluation was introduced. Relative to the teacher evaluation phase, in classroom A a mean of .8 fewer tests were passed per day, while in classroom B, a mean of .4 fewer tests were passed per day. The application of contingent self-evaluation had a much more marked effect in each classroom on

TABLE 1

Mean Performance Measures of Tests Passed, Tests Failed, and Tests Taken Per Day for Classroom A and B.

	CLASSROOM A		CLASSROOM B	
	Teacher Evaluation	Contingent Self-Evaluation	Teacher Evaluation	Contingent Self-Evaluation
Mean tests passed/day	3.0	2.2	2.4	2.0
Mean tests failed/day	1.6	.3	2.1	.7
Mean tests taken/day	4.6	2.5	4.5	2.7

the mean number of tests failed per day and the mean number of tests taken per day. Mean daily test failures decreased by an average of 1.3 per day in classroom A (a decrease of over 80 percent), while the mean number of tests taken per day decreased by 2.1. In classroom B, mean daily failures were reduced by 1.4 per day (a decrease of over 65 percent), while the mean number of tests taken per day decreased by 1.8.

The dramatic change in the daily failure rate can be seen in Figure 2, which presents cumulative daily test failures for both classrooms, throughout the investigation.[2] In classroom A the teacher evaluation phase was characterized by a relatively stable and high rate of failure. In contrast, when contingent self-evaluation was introduced, the phase was marked by several periods of nonfailure, interrupted by one or two days on which a single test was failed. In classroom B, teacher evaluation resulted in a very stable and high rate of failure. Contingent self-evaluation produced an immediate decline in failures for the first seven days. During the next thirteen days failures increased, although the rate of failure was still below that seen during teacher evaluation. During the final five days of the contingent self-evaluation phase, no tests were failed in classroom B.

Data on individual students support the group findings that contingent self-evaluation improved academic performance in both classrooms. All students in each classroom earned the opportunity to self-evaluate during the treatment phase. Six students in classroom A earned the privilege during the first week that the contingency was in effect, and all

---

[2]The phases appear to be of shorter duration in Figure 2 than previously stated. This is due to the fact that on several days, in each phase, no tests were taken.

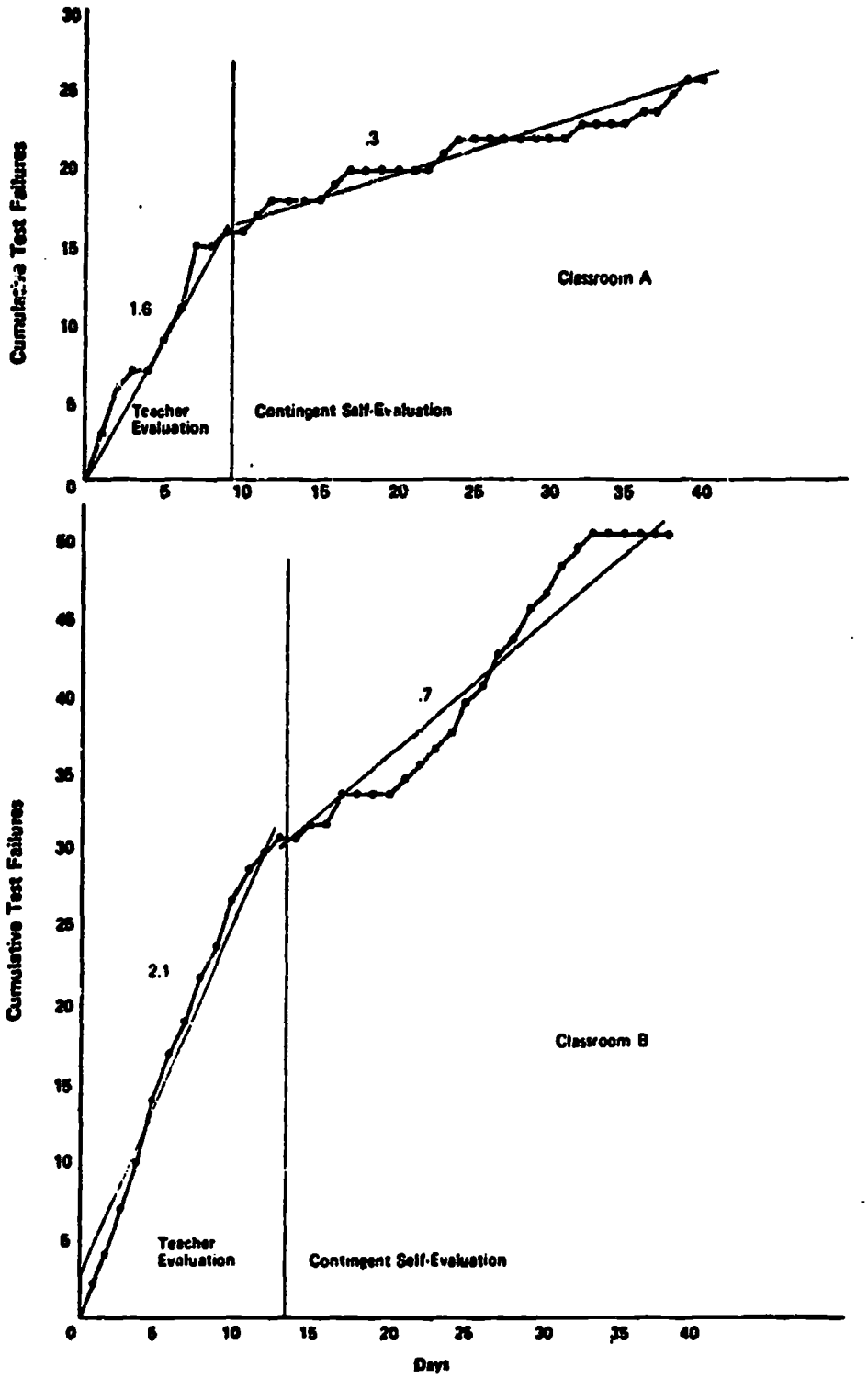

Figure 2. Daily cumulative test failures for classrooms A and B. Also shown is the mean daily failure rate for each phase.

but one child in classroom A earned the privilege by the end of the second week. Seven children in classroom A passed all remaining tests once they had earned the self-evaluation privilege, and thus continued to self-evaluate for the duration of the study.

In classroom B ten children achieved self-evaluation status during the first week, and all but one child had earned the privilege by the third week of the contingent phase. Five children passed all remaining tests once they had become self-evaluators, and were, therefore, able to retain the privilege throughout the remainder of the investigation.

Table 2 presents the number of individual children in each classroom whose test performance increased, decreased, or showed no change when contingent self-evaluation was introduced. With regard to

TABLE 2

Summary of Changes in Individual Test Performance for Children in Classrooms A and B When Contingent Self-Evaluation was Introduced.

	CLASSROOM A			CLASSROOM B		
	Increase	Decrease	No Change	Increase	Decrease	No Change
Percent of tests passed	8	2	3	9	2	1
Mean percent correct	7	5	1	8	4	---
Mean tests passed/day	4	9	---	4	6	2
Mean tests failed/day	2	7	4	1	11	---
Mean tests taken/day	1	12	---	3	8	1

percent of tests passed, seventeen children in classrooms A and B improved their performance, while only four children showed a decrease. Of the four students who exhibited no change, each passed all tests taken during the two phases of the study. In terms of mean percent correct, fifteen students in the two classrooms showed improved performance. Of the nine whose mean test scores decreased, four dropped from averages of 100 percent to an average of 92 percent or above, and all but two of the nine students maintained averages higher than the minimum passing level of 85 percent. While more children showed a decrease than an increase in tests passed per day, ten children either increased or remained the same on this variable. Thus, the group statistic, which indicated only a slight decrease in mean daily passes, is supported by the individual data.

The large majority of students showed a decrease in mean daily failures. Of the four children who exhibited no change on this variable, three of them passed all of their tests in each phase of the investigation. Finally, it can be seen that 80 percent of the students attempted fewer tests during the contingent self-evaluation phase than they had during teacher evaluation.

Discussion

The data indicate that the introduction of the contingent self-evaluation procedure produced improved academic performance for the majority of children in the two classrooms. More importantly, it was demonstrated that children could perform at a high level while sharing the responsibility for managing their own behavior. It appears that students are more able to successfully control their own learning experiences than was previously believed.

The contingent self-evaluation phase, when compared to the teacher evaluation condition, produced five distinct changes in test performance in the two classrooms. The percent of tests passed increased, the mean score on each test increased, mean daily test failures and mean daily tests attempted decreased markedly, and the mean number of tests passed daily decreased slightly. Because most children continued to pass tests at a fairly high rate, relative to tests attempted and failed, it would appear that the students began attempting tests only when they were confident of passing them.

This possibility, if true, would have implications for the management of these and similarly designed classrooms. In the teacher evaluation condition, children were assigned tests at the point at which they met the 100 percent accuracy criterion on all workbook pages in a given skill. It was generally assumed that mastery of these pages implied test "readiness." However, in the contingent self-evaluation phase students were free to determine when to attempt a test. They responded with a lower rate of test-taking, thus possibly indicating a need for additional review before attempting a test. It is also possible that for those students who actually engaged in this review or additional study behavior, that the review itself, rather than the contingency, may have maintained high test performance. The contingency, for some students, may have just produced the initial change in test behavior. For other students, however, who did not use the evaluation time to review work, the contingency may have been the major factor in maintaining high test performance. Another possibility is that some students may have found daily teacher evaluation to be an aversive condition. During the initial phase of the study students did not have the option of avoiding teacher assessment. However, when given the opportunity to avoid teacher corrections during the self-evaluation stage, by improving their test performance, they worked to increase their test accuracy.

The present study did not attempt to determine which component of the self-evaluation procedure produced the observed changes. Several investigations need to be conducted to examine this and related questions. A study now in progress is examining both contingent and noncontingent self-evaluation procedures to determine if a contingency is required to produce improved performance, or whether unearned self-evaluation is sufficient. Future investigations will also attempt to look more closely at student behavior during self-evaluation in an effort to determine how students use the evaluation time.

One point should be made about the type of curriculum material used in this study and its relationship to the validity of the results. It has been noted by Brigham, Finfrock, Breunig, and Bushell (1972) that there are distinct advantages in using an individualized, programmed curriculum when the dependent variable in an investigation is academic performance. They state that programmed materials provide small steps of equivalent difficulty and the opportunity for self-pacing. The first factor enables realistic comparisons to be made between student performance at different points in time. The second factor results in students being spread throughout the curriculum and, hence, adds to the credibility of the independent variable in accounting for any observed changes. The present investigation, therefore, adds to the small body of classroom-management research in which individualized materials have been used. It is hoped that the points made by Brigham et al., are seriously considered when other investigators undertake research in this area.

Finally, it should be noted that the teachers and the investigators would have been satisfied if the average level of test performance obtained during the contingent self-evaluation phase had equaled, rather

than exceeded, the level observed during teacher evaluation. Had that been the case, the investigation would still have been viewed as a success, because some students would have been managing their own behavior, and overall performance would have been maintained.

Hopefully, future research in this area will result in procedures that will enable students to gain additional control over their own educational process.

References

Brigham, T. A., Finfrock, S. R., Breunig, M. K., & Bushell, D., Jr. The use of programmed materials in the analysis of academic contingencies. Journal of Applied Behavior Analysis, 1972, 5(2), 177-182.

Broden, M., Hall, R. V., & Mitts, B. The effect of self-recording on the classroom behavior of two eighth-grade students. Journal of Applied Behavior Analysis, 1971, 4(3), 191-199.

Glynn, E. L. lassroom applications of self-determined reinforcement. Journal of Applied Behavior Analysis, 1970, 3, 123-132.

Glynn, E. L., Thomas, J. D., & Shee, S. M. Behavioral self-control of on-task behavior in an elementary classroom. Journal of Applied Behavior Analysis, 1973, 6(1), 105-113.

Kaufman, K., & O'Leary, K. D. Reward, cost and self-evaluation procedures for disruptive adolescents in a psychiatric hospital school. Journal of Applied Behavior Analysis, 1972, 5, 293-309.

Knapczyk, D. R., & Livingston, G. Self-recording and student teacher supervision: Variables within a token economy structure. Journal of Applied Behavior Analysis, 1973, 6(3), 481-486.

Lovitt, T., & Curtiss, K. Academic response rate as a function of teacher and self-imposed contingencies. Journal of Applied Behavior Analysis, 1969, 2, 49-54.

Santogrossi, D. A., O'Leary, K. D., Romanczyk, R. G., & Kaufman, K. F. Self-evaluation by adolescents in a psychiatric hospital school token program. Journal of Applied Behavior Analysis, 1973, 6(2), 227-287.