DOCUMENT RESUME

ED 083 134 SP 006 890

AUTHOR Fox, Karen F. A.
TITLE Incentives for School Achievement.
PUB DATE Feb 73
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, February 1973

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Incentive Systems; Learning; *Motivation; Parent Influence; *Parent Participation; Reading Research; *Reading Skills
IDENTIFIERS Jung (S M)

ABSTRACT

In 1971, the Wethersfield, Connecticut School Department received a U.S. Office of Education grant to initiate a study testing some of the incentive models proposed by Jung and his colleagues. The goal of this project was to improve students' initial mastery of reading skill objectives by a program of assigning appropriate objectives to each student, testing mastery of assigned objectives through criterion-referenced tests, and finally awarding incentives to target groups based on students' mastery of assigned objectives. In addition, the project has the goal of implementing and evaluating the impact of three incentives. Four experimental groups were devised: Student, Parent, Parent-Student, and Control (no treatment). In the Student Incentive Treatment, students received instruction, Regional Center Materials, and the 1972 10X Catalog of toys, books, or records, and, ultimately, a class field trip. Incentives in the Parent Incentive Treatment included parents praising student success as an incentive with the parents receiving as an incentive $200 if 80 percent of the children in the class achieved 80 percent of their assigned objectives with 80 percent accuracy. The Parent-Student Treatment combined the incentives of the first two groups. (The document contains information on the analysis of the reading achievement outcomes and classroom observation data, supportive tables, and a discussion of the results.) (For related document, see ED 062 668.) (JA)

INCENTIVES FOR SCHOOL ACHIEVEMENT

Karen F. A. Fox

American Institutes for Research

Palo Alto, California

Paper presented at the Annual Meetings of the American Educational Research Association, New Orleans, 28 February 1973.

INCENTIVES FOR SCHOOL ACHIEVEMENT

Karen F. A. Fox,[1] American Institutes for Research

A recent article on performance contracting (Elam, 1970) reported the following statement made by a teacher: "I would stand on my head if that would make them learn, but I wonder if the children are really learning or just storing a little knowledge for long enough to get the reward." The teacher's point has received increasing attention--What is the effect of using incentives to influence student achievement? Ample evidence from the experimental analysis of human behavior (e.g., Ulrich et al., 1966) suggests that judicious applications of reinforcements can effectively alter the frequency and form of children's classroom behavior. In fact, the bulk of studies of reinforcement in school settings has focused on changing various kinds of disruptive behavior through the use of small reinforcers delivered at frequent intervals. The influence of reinforcement techniques on measured achievement in school subjects is by no means as well demonstrated, although preliminary studies by Jung and Lipe (1971) demonstrate the feasibility of implementing such incentive systems in educational settings.

Jung, Lipe, and Wolfe (1971) made an extensive review of the use of a wide range of reinforcers in educational settings, including knowledge of results, social reinforcers, aversive stimuli, secondary reinforcers such as tokens, and material incentives. In addition they examined several types of delivery systems (e.g., token economics, performance contracting, etc.), timing of reinforcements, and possible target populations who might receive

---

[1]The author is now on the faculty of Teachers College, Columbia University.

incentives, including students, teachers, parents, and others.

In 1971 the Wethersfield, Connecticut School Department received a U.S. Office of Education grant to initiate a study testing some of the incentive models proposed by Jung and his colleagues. While incentives are typically awarded at frequent intervals for small observable increments in student behavior, the decision was made to award students substantial material incentives for relatively large units of performance (such as progress over a period of several months), and to give other types of reinforcers at shorter intervals.

Theories of reinforcement suggest that a substantial time lapse between the reception of the incentive and the learning upon which the incentive was contingent might cause the incentive itself to have rather little impact on the students' effort to improve their performance, especially since improvement over time requires sustained, cumulative learning. Nevertheless, macroincentives (substantial material incentives for relatively large units of behavior) seem highly likely to give good results, if certain conditions are met:

(1) The material incentives are made contingent upon some well-publicized gain in level of performance.

(2) The most functional and desirable incentives are selected.

(3) Student group involvement is fostered by offering some incentives to entire classes based on group performance.

(4) Teachers are encouraged to offer extra help and tutoring to those students who desire it.

(5) Intervals shorter than a full school year are identified to allow for more frequent incentive delivery.

The goal of the Wethersfield Incentives Project was to improve the reading skills of students in grades 1 through 4 by means of a program of assessing students' initial mastery of reading skill objectives, assigning appropriate objectives to each student, testing mastery of assigned objectives through criterion-referenced tests, and finally awarding incentives to target groups based on students' mastery of assigned objectives. In addition, the Project had the goal of implementing and evaluating the impact of three incentive treatments.

Experimental Design

To permit evaluation of the effects of various incentive treatments on student reading achievement, a four-by-four experimental research design was initially selected, with one teacher and her/his classroom in each cell. Four teachers were randomly selected from the pool of all teachers at each of four grade levels in nine of the town's elementary schools (including one parochial school). The teachers at each grade level were randomly assigned to one of four treatments: Student, Parent, Parent-Student and Control. The selected teachers attended a two-week workshop during which they received training in preparing performance objectives. During the second week the teachers met by grade level and prepared reading objectives which they wanted 80 per cent of their students to be able to accomplish during the school year. The teachers also received instruction in principles of

reinforcement, writing criterion-referenced tests, the nature of the Incentives Project, and the particulars of the treatments which they would be carrying out.

There was some concern that teachers assigned to the Control (no treatment) group might be influenced, by their participation in the summer staff development workshops, to teach differently than a "true" control with no knowledge of the design and purposes of the study. Moreover, when placed in competition with an experimental group using an innovative procedure which threatens to replace the control (usually traditional) procedure, a control group may exhibit above-average performance--what Saretsky (1972) describes as the John Henry effect. To allow further comparisons, the evaluator and the Project administrative staff selected two classrooms at each grade level as a Passive Control group. Students in those eight classes took pre and post-project standardized tests, but the students and their teachers were not informed about the Project and were not included in Project activities or incentives. The initial Control group was designated as the Active Control group.

Description of Incentive Treatments

Three incentive treatments were implemented. All three treatments shared the following elements:

(1) The focus was on improving the reading achievement of students.

(2) Teachers received two weeks of intensive training in writing performance objectives and criterion-referenced tests, and in the purposes and methods of the Project.

(3) Students' entry-level reading skills were determined by an A.I.R.-constructed criterion-referenced test. Teachers selected and assigned appropriate objectives individually to their students, keeping records of assigned and mastered objectives for each student.

(4) There were four incentive earning periods, each lasting approximately six weeks.

(5) Experimental groups received incentives contingent upon students' successful mastery of assigned reading objectives, as measured by criterion-referenced tests.

(6) A criterion-referenced test developed by the Assistant Project Director was administered at the end of the second incentive earning period, and served as the basis for awarding incentives. Results of a second A.I.R.-constructed criterion-referenced test determined incentive awards for the final incentive period.

(7) Standardized reading tests were administered as pre and post measures but were not used in determining incentive awards.

The Student incentive treatment. Students received incentives, purchased out of Project funds, at the end of each of the four incentive earning periods. At the end of the first and third earning periods, each student individually selected and received a toy, book, or record (costing $5.00 to $10.00) if she/he had demonstrated mastery on all of her/his assigned objectives by correctly answering 80% of the items on each objective on a teacher-constructed criterion-referenced test. At the end of the second

and fourth earning periods, the entire classroom shared in a group incentive if 80 percent of the students achieved 80 percent of their assigned objectives with 80 percent accuracy. Virtually all the classes selected field trips for group incentives.

The Parent incentive treatment. At fall orientation meetings parents were asked to tutor and to reinforce their child in order to enhance the child's motivation and achievement in reading. Parents were told that their children would bring home "blue slips" from their teachers each time they completed an important assigned objective. The "blue slip" told the parents, "[child's name] has shown accomplishment of the objectives set for him. Please show how pleased you are by (1) Praising your child, and (2) Offering him a special treat of his own choosing." In addition to being asked to reinforce their child, parents were told that teachers might occasionally send home "parent tutoring forms," giving a specific objective on which their child needed more practice, and suggesting materials and techniques for home tutoring. At the end of the second and fourth incentive periods, the parents in each classroom were given $200 if 80 percent of the children in the class achieved 80 percent of their assigned objectives with 80 percent accuracy. The parents most frequently selected a field trip for the students in the class, or some activity which both they and their children could share.

The Parent-Student incentive treatment. This treatment was a combination of the Parent and the Student treatments--in effect, a "double dose" of

reinforcement. Students received individual and group incentives purchased from project funds, while parents as a group received $200 per classroom twice a year, and determined how to spend these incentives.

Control groups. Students in the Active and Passive Control groups did not receive incentives, and teachers were expected to teach their classes as they normally would.

Reading Achievement Outcomes

Determination of the impact of the incentive treatments as made in three ways:

(1) Students' performance on selected standardized tests;

(2) Whether children mastered their assigned objectives and thus earned incentives; and

(3) Students' performance on criterion-referenced pre- and posttests of reading skills.

Table 1 lists the instruments which were used in the Project.

Insert Table 1 about here

The initial data analysis plan was to use analysis of covariance (with standardized test scores as the covariate) to determine whether the three incentive treatment groups had gained significantly more during the Project

as measured by the selected standardized tests than did the Active and Passive Control groups. Preliminary use of analysis of covariance on the data indicated, however, that its use was not statistically justified, since the effects of the various experimental treatments were not additive.

Instead, significance tests were performed to determine whether the experimental and Active Control groups made gains significantly greater than would be expected in the absence of treatment. Results on the California Achievement Test are presented in Table 2. The mean gain of the Passive Control group was used as an expectation standard, i.e., as the expected gain in the absence of treatment. The mean gain of the total Passive Control group, 1.81 points, was subtracted from the mean gain of each of the three experimental groups and the Active Control group, given in the rows designated "Total group," yielding the results in the column "Gain above Expected Gain." As can be seen in the "Gain above Expected Gain" column, every treatment group including the Active Controls exhibited greater gains than did the comparable Passive Control group. The significance tests are used to determine if the students gained significantly more than their comparable Passive Control group. In most cases, they did.

Insert Table 2 about here

Standardized tests were also administered to first graders. However, since the test data did not meet the assumptions for using analysis of covariance,

and since the tests administered to first graders were distinct from those administered to students in grades 2 through 4, the data for first graders was not extensively analyzed.

Frequently, it is suggested that incentives will have a much greater impact on the motivation and achievement of slower students than of capable students who are already being rewarded by their success in school tasks. In order to examine this hypothesis, the students in each classroom were also ranked on the basis of their pretest raw scores on the California Achievement Test. The top one-third of the students in each classroom was combined within the five treatments, and the bottom one-third was likewise combined. As in the analysis of the toal treatment groups, the average gains of students in the top and bottom thirds of the Passive Control group were used as expectation standards. Since the Passive Control group classes were randomly selected and received only regular classroom instruction, it is reasonable to assume that students in the treatment classrooms would have made equivalent gains in the absence of the treatment.

While the top third achieved significant gains only in the Parent group, the bottom third made significant gains in the four experimental and Active Control groups. The total treatment groups gains above expected gains were also significant at the one percent level.

Criterion-referenced tests were constructed by American Institutes for Research to correspond directly to the student instruction product objectives written by participating teachers. Mastery of each objective was assessed

by one or more items. The results of the criterion-referenced tests served two purposes. First, each student's test was scored to determine which of the individually assigned objectives she/he had mastered. If, at the end of the second and fourth earning periods, 80 percent of the students in the Student and Parent-Student treatment classrooms achieved 80 percent mastery on the criterion-referenced tests, the entire classroom shared in a group incentive. All eight classrooms accomplished this goal at the end of the second earning period, and all but one classroom did so at the end of the fourth period as well. Individual incentives were awarded at the end of the first and third earning periods based on individual mastery of assigned objectives determined by teacher-made criterion-referenced tests. All but one student earned individual incentives at the end of the first period, and all earned incentives for the third period.

Second, the results were analyzed to permit comparisons between treatment groups and are presented in Table 3. Tests were scored by assigning two points for 100 percent mastery of an objective; one point for partial mastery --over 80 percent but less than 100 percent accuracy; and zero points for less than 80 percent mastery. Since there were 130 reading objectives, the maximum score a student could receive was 260 points, by attaining 100 percent mastery on all objectives. An interim criterion-referenced test written by the Assistant Project Director was administered half-way through the Project solely to provide a basis for awarding group incentives. The results of that test are not included as part of the evaluation.

Insert Table 3 about here

Table 3 presents the criterion-referenced test results by top and bottom third and by total group for grades 2, 3, and 4 combined. The top and bottom thirds include the same students that are included in the breakdowns in Table 2. The numbers of students vary slightly from Table 2 due to the absence of criterion-referenced test data for some students. First graders were not included in the analyses in Table 3, since they did not take the California Achievement Test upon which the breakdowns were made. Since the criterion-referenced reading tests were not administered to the Passive Control group, there was no way to establish an expectation standard for use as a basis for significance tests of gains. The Passive Control group was not given the criterion-referenced tests since to have done so would have alerted both teachers and students that they were actually part of an experimental study. (The California Achievement Test was routinely given all students and therefore did not constitute such a signal.) As Table 3 indicates, the top and bottom thirds and the total group of the Active Control group all achieved greater point gains than did all but the top third and total group in the Parent treatment. Possible explanations of these results will be discussed later.

Classroom Observations

Incentives alone may provide motivation, but they do not provide the means

for improving student achievement. In the Wethersfield Incentives Project teachers were trained in selecting and preparing behavioral objectives, in assigning appropriate objectives to individual students, and in assessing mastery of assigned objectives. As a result of the Project emphasis on mastery of objectives and individualization based on criterion-referencing testing, we would expect to see some differences between the experimental classrooms and the control classrooms. Specifically, the following hypotheses were proposed:

1. Experimental teachers will use praise and reward stipulation more frequently than will Active Control teachers.
2. Active Control teachers will use punishment and punishment stipulation more frequently than will experimental teachers.
3. Experimental students will spend more time in individual and small group interaction with the teacher during reading than will Active Control students.
4. Experimental students will spend more time studying alone or with peers than will Active Control students.
5. Experimental students will exhibit less non-productive behavior than will Active Control students.

The classroom observation data was collected over ten observation periods of two to three weeks each by seven trained volunteer observers. During each observation period each class was observed for a total of forth minutes, twenty minutes of reading and twenty minutes of some other subject.

To test the five hypotheses the observations for each of the five relevant categories were totally for all ten observation periods, by experimental and control (Active Control) treatments. The means and percentages of total observed class time for both groups for the five categories were virtually identical, as shown in Table 4, indicating virtually no significant differences between experimental and Active Control classroom activity patterns when we look at the combined data for the entire course of the Project. (The category of criticizing or threatening punishment incorporates such a relatively small percentage of total observed time, that while the difference between experimental and control classes is highly significant statistically, the difference is an artifact of a drastic but temporary rise in criticism/punishment by control teachers during two observation periods in the middle of the year.)

Insert Table 4 about here

In order to examine possible time shifts in behavior categories over time, the ten observation periods were collapsed into five time periods, and mean percentages of reading class time spent in various activities were graphed for experimental and Active Control classes. As the similarity of means for the year would suggest, there were no clear trends favoring experimental over Active Control classes, and the five hypotheses, with the possible exception of Hypothesis 2, were not supported.

Discussion of Results

When a controlled laboratory study yields inconclusive results, the range of explanatory factors tends to be somewhat easier to grasp and scrutinize than in a school-based study in which virtually all phases of the experimental treatment are mediated by human beings with varying degrees of insight into and commitment to the goals of the project. While this conclusion is tautological, it may be instructive to describe the pitfalls of implementing a large-scale incentive system. Particularly when the efficacy of several treatments is being compared (versus treatment/no treatment), the treatments must be rigorously and uniformly carried out. In this study teachers varied in how they assigned reading objectives and in how they awarded incentives.

- One teacher assigned all the objectives to each child, even though the child had already demonstrated mastery of at least some of the objectives on the pretest. The teacher felt that doing the work again would serve to "fix" it in the child's mind.
- One teacher viewed working on the objectives as "extra work," to be done only after other school work was done.

Students in the Parent and Parent-Student classes were to receive "blue slips" to take home to their parents when they had mastered a key objective, one which would be of an appropriate difficulty level so that the child would be able to master it in two weeks. In the absence of specific criteria, two teachers sent home blue slips for mastery of each objective

and subobjective. One teacher sent home one blue slip per week ("every Friday"), but required faster students to accomplish more objectives than slower ones. Two teachers gave one blue slip per earning period (6 to 8 weeks) per child. One based awards of blue slips on the student's attitude, behavior in class, effort--and how well he did on his objectives. Another teacher gave very few blue slips, stating that "phone calls to parents are more personal." During the second semester the number of blue slips sent by each teacher varied from none to 175, with a mean of 40 slips for eight teachers. Since particularly in the Parent group the students were rewarded primarily by their parents for bringing home blue slips, many children received very few (or no) incentives because they were given few (or no) blue slips. To the extent that a blue slip represents a "unit of value" as a reinforcer that has approximately the same value for all recipients, it is clear that in effect many children did not receive much reinforcement.

Parental reactions to their role as dispensers of reinforcement varied considerably. Few parents attended orientation meetings in the fall, where the parental role was explained. As a result many parents equated "incentives" with "bribes," a perspective which was confirmed by some children who became hustlers--presenting blue slips to their parents, and immediately demanding, "What are you gonna give me?" Parents who saw their children passed over by the conventional reward structure of the school (e.g., grades, making rapid progress) generally favored incentives (material and

social) when they saw it contribute to the motivation and satisfaction of their children. But several parents whose children were in the Parent group (and thus did not receive material incentives at school) were miffed that they had to pay for incentives for their children, while other children got rewards purchased from Federal funds.

A potential means of improving the orientation and training of teachers (and perhaps of parents) participating in incentive programs is a multimedia package developed in late 1972 under the Targeted Communication Program of U. S. O. E. (Lipe, Weisgerber, and Fox, 1972). The package includes a two-part slide-tape introduction to principles of behavior modification and implementation of incentive programs, ranging from programs for individual students to large-scale programs such as that used in Wethersfield. A workbook parallels the slide-tape presentations, and a resource guide details the principal topics and suggests additional sources of information. The printed portions of the package should be available soon through the ERIC system.

The Wethersfield Incentives Project is now in its second year, with some restructuring of treatments, training, and evaluation procedures, and with more students and grade levels involved, allowing for comparisons between students who have been in the Project from its inception in Fall 1971 and those who entered the Project in Fall 1972.

REFERENCES

Elam, S. M. The age of accountability dawns in Texarkana. Phi Delta Kappan, June 1970, pp. 509-514.

Fox, K. F. A. Evaluation of the Wethersfield incentives project. Palo Alto, Calif.: American Institutes for Research, 1972.

Jung, S. M., Lipe, D. H., & Canter, S. Experimental assessment of an incentive program to enhance school learning: a pilot study. Palo Alto, Calif.: American Institutes for Research, 1971.

Jung, S. M., Lipe, D. H., & Wolfe, P. S. Study of the use of incentives in education and the feasibility of field experiments in school systems. Palo Alto, Calif.: American Institutes for Research, 1971.

Saretsky, G. The OEO P. C. experiment and the John Henry effect. Phi Delta Kappan, May 1972, pp. 579-581.

Ulrich, R., Stachnik, T., & Mabry, J. (eds). Control of human behavior. Glenview, Ill.: Scott, Foresman, 1966.

TABLE 1

Summary of the Evaluation Instruments
for the Reading Component

Test	Administration			
	Pretest	Posttest	Treatment Groups	Grade Levels
A.I.R. criterion-referenced reading tests	X (Form A)	X (Form B)	Experimental and Active Control	1-4
California Short Form Test of Mental Maturity	X		Experimental and Active Control	2-4
California Achievement Tests, Level 2	X (Form A)	X (Form B)	Experimental, Active and Passive Control	2-4
Metropolitan Readiness Test	X	X	Experimental	1
Primary Mental Abilities Test	X		Experimental	1

TABLE 2

Student Gains in Total Raw Score on the California Achievement Test for Grades 2, 3, and 4 Combined

Treatment Group	N	Mean Pretest CAT Total Score (Form A)	Mean Posttest CAT Total Scare (Form B)	Mean Gain	Gain above Expected Goin*	t Value	Significant
Student							
Top third	22	66.82	68.59	1.77	1.96	.99	No
Bottam third	22	49.95	61.09	11.14	7.64	4.19	Yes**
Total group	69	59.20	66.41	7.21	5.40	4.26	Yes**
Parent							
Top third	20	69.30	74.25	4.95	5.14	2.53	Yes**
Bottom third	20	46.15	54.90	8.75	5.25	1.96	Yes***
Total group	61	57.79	64.72	6.93	5.12	3.71	Yes**
Parent-Student							
Top third	22	76.05	78.14	2.09	2.29	1.22	No
Bottom third	22	52.86	67.18	14.32	10.82	3.39	Yes**
Total group	65	64.78	72.96	8.18	6.37	3.94	Yes**
Active Cantral							
Top third	17	73.53	75.35	1.82	2.01	1.20	No
Bottam third	17	49.37	57.84	8.47	4.97	2.37	Yes***
Total group	53	60.55	66.12	5.57	3.76	3.00	Yes**
Passive Cantrol							
Top third	42	75.69	75.50	-.19	0.00		
Bottom third	42	59.19	62.69	3.50	0.00		
Total group	124	68.31	70.12	1.81	0.00		

* Gain above Expected Gain for each group is the mean gain of the group minus the mean gain of the comparable Passive Control group.

** Significant at the 1% level.

*** Significant at the 5% level.

TABLE 3

Student Gains in Points on the Criterion-Referenced Tests for Grades 2, 3 and 4 Combined

Treatment Group	N	Mean Pretest CR Total Score (Form A)	Mean Pretest CR Total Score (Form B)	Mean Gain
Student				
Top third	22	143.91	170.00	26.09
Bottom third	20	115.15	149.45	34.30
Total group	67	129.10	161.13	32.03
Parent				
Top third	17	147.06	195.59	48.53
Bottom third	18	100.22	143.78	43.56
Total group	54	120.96	166.35	45.39
Parent-Student				
Top third	21	158.05	195.71	37.67
Bottom third	20	130.10	168.60	38.50
Total group	62	146.52	183.58	37.06
Active Control				
Top third	15	139.87	185.73	45.87
Bottom third	17	94.76	139.76	45.00
Total group	48	114.67	164.12	49.46

TABLE 4

Means and Percentages of Total Observed Reading Class Time Spent in Selected Behavior Categories for Experimental Classes and Active Control Classes

	Experimental Classes		Active Control Classes	
Behavior Category	Mean	Percentage of Total Time	Mean	Percentage of Total Time
Student				
Studying alone or interacting with peer(s)	55.86	23.28%	56.69	23.62%
Interacting with the teacher individually or in small group	26.09	10.87	26.06	10.86
Non-productive behavior (obstrusive or unobtrusive)	16.37	6.82	14.36	5.98
Teacher				
Praising a student or offering a performance-contingent reward	4.35	1.81	3.79	1.58
Criticizing a student or threatening punishment	1.02	.43	2.08	.87