INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms
300 North Zeob Road
Ann Arbor, Michigan 48106
A Xerox Education Company
GREGORY, Lynne Ann, 1946-
THE RELATIVE EFFECTIVENESS OF POSITIVE REINFORCEMENT AND RESPONSE COST PROCEDURES IN A TOKEN REINFORCEMENT PROGRAM IN TWO SPECIAL EDUCATION CLASSES IN A JUNIOR HIGH SCHOOL.

The Ohio State University, Ph.D., 1972
Psychology, clinical

University Microfilms, A XEROX Company, Ann Arbor, Michigan
THE RELATIVE EFFECTIVENESS OF POSITIVE REINFORCEMENT
AND RESPONSE COST PROCEDURES IN A TOKEN REINFORCEMENT
PROGRAM IN TWO SPECIAL EDUCATION CLASSES
IN A JUNIOR HIGH SCHOOL

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Lynne Ann Gregory, B.A., M.A.

The Ohio State University
1972

Approved by

[Signature]
Adviser
Department of Psychology
PLEASE NOTE:

Some pages may have
indistinct print.
Filmed as received.

University Microfilms, A Xerox Education Company
ACKNOWLEDGMENTS

As with any research project, the author of the paper is only the most visible member of a group of contributors. Dr. Henry Leland, throughout my term as graduate student, has contributed immeasurably to the growth of my thinking in the area of exceptionality and especially has helped to illuminate the problems of mentally retarded individuals. Dr. Denny Nolan and Dr. David Hothersall passed my notions through the behaviorist's sieve and ultimately helped to strengthen the research design; if faults are apparent, it is because I was unable to implement their suggestions. Dr. Malcolm Helper contributed much time in reading several drafts of the paper and helped to clarify my thoughts on several important issues.

Marvin Eisen shared in the project from its conception; he served as mentor, gadfly, and editor through all the periods of crisis and elation and was truly my backup reinforcer. Michael Allen is to be commended for responding promptly and efficiently to all my calls of distress regarding the problems of doing a dissertation from afar.

My observer, Jean Anne Matter, accepted a cut in salary to help in the project, and she should be thanked for all the early mornings when she was on time and alert and for her suggestions as the investigation unfolded.
Finally, much appreciation is due the teachers and principal in the Pomona Unified School District who allowed me to invade their classrooms daily for over three months and who accepted my endless suggestions on how to engineer the teaching process.
VITA

February 8, 1946 . . . . Born - Longview, Texas

1968 . . . . . . . . . B.A., Louisiana State University, Baton Rouge, Louisiana

1968-1969. . . . . . . . University Fellow, Department of Psychology, The Ohio State University, Columbus, Ohio

1969-1970. . . . . . . . Teaching Associate, Department of Psychology, The Ohio State University, Columbus, Ohio

1970-1971. . . . . . . . Psychology Trainee, Nisonger Center for Mental Retardation, The Ohio State University, Columbus, Ohio

1971-1972. . . . . . . . Dissertation Fellow, Department of Psychology, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Developmental Psychology

Studies in Development. Professors George G. Thompson, John E. Horrocks, Malcolm M. Helper

Studies in Testing. Professors Donald C. Smith, Donald Bersoff

Minor Field: Mental Retardation

Studies in Exceptionality. Professors Henry Leland, Reginald C. Jones
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background to the Problem and Previous Research</td>
<td></td>
</tr>
<tr>
<td>Aims of the Present Research</td>
<td></td>
</tr>
<tr>
<td>II. METHOD</td>
<td>27</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Setting</td>
<td></td>
</tr>
<tr>
<td>Observation Technique</td>
<td></td>
</tr>
<tr>
<td>Behavior Definitions</td>
<td></td>
</tr>
<tr>
<td>Experimental Conditions</td>
<td></td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>42</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
</tr>
<tr>
<td>Classroom Behavior of EMR Students</td>
<td></td>
</tr>
<tr>
<td>With Token Reinforcement</td>
<td></td>
</tr>
<tr>
<td>Comparison of Adaptive Behavior</td>
<td></td>
</tr>
<tr>
<td>Levels in Normal and EMR Classrooms</td>
<td></td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>72</td>
</tr>
<tr>
<td>V. SUMMARY</td>
<td>81</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>83</td>
</tr>
<tr>
<td>B</td>
<td>84</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>85</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Average Percentages for Specific Behavior Classes for Each Experimental Phase in Ms. B's Class ........................................... 49

2. Average Percentages for Specific Behavior Classes for Each Experimental Phase in Study Hall .................................................. 53

3. Average Percentages for Specific Behavior Classes for Each Experimental Phase in Ms. A's Class .................................................. 64

4. Average Percentages for Specific Behavior Classes for Normal and EMR Classes in the Baseline Condition ........................................ 69
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for the Subjects in Ms. B's Class under each Treatment Condition.</td>
<td>44</td>
</tr>
<tr>
<td>2.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S1 and S2 in Ms. B's Class under Each Treatment Condition.</td>
<td>46</td>
</tr>
<tr>
<td>3.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S3 and S4 in Ms. B's Class under Each Treatment Condition.</td>
<td>47</td>
</tr>
<tr>
<td>4.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S5 and S6 in Ms. B's Class under Each Treatment Condition.</td>
<td>48</td>
</tr>
<tr>
<td>5.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for the Students in the Study Hall under Each Treatment Condition.</td>
<td>51</td>
</tr>
<tr>
<td>6.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for the Subjects in Ms. A's Class under Each Treatment Condition.</td>
<td>55</td>
</tr>
<tr>
<td>7.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S3 and S4 in Ms. A's Class under Each Treatment Condition.</td>
<td>57</td>
</tr>
<tr>
<td>8.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S5 and S7 in Ms. A's Class under Each Treatment Condition.</td>
<td>58</td>
</tr>
<tr>
<td>9.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S8 and S9 in Ms. A's Class under Each Treatment Condition.</td>
<td>59</td>
</tr>
<tr>
<td>10.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S10 and S11 in Ms. A's Class under Each Treatment Condition.</td>
<td>60</td>
</tr>
<tr>
<td>11.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S12 and S13 in Ms. A's Class under Each Treatment Condition.</td>
<td>61</td>
</tr>
<tr>
<td>12.</td>
<td>Mean Percent of Intervals of Appropriate Behavior for S14 and S15 in Ms. A's Class under Each Treatment Condition.</td>
<td>62</td>
</tr>
</tbody>
</table>
13. Mean Percent of Intervals of Appropriate Behavior in the EMR and Normal Classrooms during the Baseline Condition .......... 66

14. Adjusted Mean Percent of Intervals of Appropriate Behavior in the EMR and Normal Classrooms during the Baseline Condition ........................................ 67

15. Mean Percent of Intervals of Appropriate Behavior in EMR Classrooms with Token Reinforcement and a Normal Class in the Baseline Condition. .................. 70
Chapter I
INTRODUCTION

In the last decade researchers have focused a great deal of attention on the application of learning principles, especially operant conditioning techniques, to the modification of a wide range of maladaptive behaviors (e.g., Bandura, 1969; Gardner, 1971; Peterson, 1968; and Ullmann and Krasner, 1965). This orientation has been extended to the field of education where there has been much interest in the development of techniques derived from behavioral principles for use in classrooms to improve academic performance and social behavior (e.g., Hall, Lund, and Jackson, 1968; Lindsley, 1964; and Skinner, 1968).

Token reinforcement programs, one of several operant conditioning schemes available, are beginning to receive considerable attention (cf. the review by O'Leary and Drabman, 1971) because of their apparent effectiveness in improving the academic and social skills of students who do not seem to respond to ordinary classroom reinforcers such as grades and teacher approval. These so-called token "economies" have been developed in such diverse classes and groups of children and adolescents as those labeled emotionally handicapped (e.g., O'Leary and Becker, 1967), disadvantaged (e.g., Herman and
and Tramontana, 1971), delinquent (e.g., Tyler and Brown, 1968), and retarded (e.g., Birnbrauer, Wolf, Kidder, and Tague, 1965). Almost all of the reported studies have obtained some degree of success in reducing general classroom levels of disruptive behavior and in augmenting the percentage of time spent in study behavior, although there has been much variability in the effectiveness of the token programs across individual students.

While the efficacy of token reinforcement programs of one variant or another in modifying maladaptive classroom behaviors of diverse subject populations and ages has been demonstrated quite adequately, few studies compared systematically the effects of differing token economy schemes with the same subjects; and no studies have employed educable mentally retarded (EMR) children in this sort of comparative effort.

This last point is of particular significance, since the EMR children who are ultimately relegated to special education classes in the public schools are those who exhibit high frequencies of inappropriate social behavior, when compared with classes of so-called "normal" students (Kirk, 1964).

Thus, the present study was designed as an attempt to examine the relative effectiveness of two major token contingency arrangements—positive reinforcement and response costs—in modifying the disruptive social behavior
of mildly retarded adolescents in two junior high classrooms. Additionally, its purpose was to assess the relative effectiveness of both of these token schemes in building up and maintaining a level of appropriate social behavior comparable to that which could be expected from students in a "normal" classroom in the same school.

Background to the Problem and Previous Research

The proliferation of studies showing the effectiveness of behavior modification strategies in changing classroom behavior within the last five years has been especially noteworthy. For example, studies have been conducted with single subjects (Ramp, Ulrich, and Dulaney, 1971), special problem students (Broden, Bruce, Mitchell, Carter, and Hall, 1970), and entire classes (Barrish, Saunders, and Wolf, 1969). Target problem behaviors have included disruptive behavior (Hall, et al., 1968), nonattending (Packard, 1970), and encopresis (Pedrini and Pedrini, 1970). Teachers (Hall, Fox, Willard, Goldsmith, Emerson, Owen, Davis, and Porcia, 1971), other students as behavioral engineers (Surratt, Ulrich, and Hawkins, 1969), and parents (Hall, Axelrod, Tyler, Grief, Jones, and Robertson, 1972) have been used as the mediators of the reinforcement program.

Derived from learning principles, especially the operant conditioning model of Skinner and his associates
behavior modification strategies have assuming as a basic tenet that behavior, while mediated by certain constitutional factors (Bijou, 1963), is controlled by its consequences in the environment. Behavior which produces positive outcomes (positive reinforcement) or which removes aversive ones (negative reinforcement) is strengthened and more likely to occur in a similar situation in the future. On the other hand, behavior which results in unpleasant consequences (punishment) or the removal of rewarding ones (response cost) will occur less frequently. Behavior modification strategies for change entail a rearrangement of the contingencies (i.e., reliable relationships between behavior and its consequences) of reinforcement so that desirable behavior patterns are reinforced while inappropriate behaviors are punished or allowed to extinguish.

The behavioral approach makes no a priori assumptions about inferred deficits of the subject population (e.g., brain damaged, retarded, or emotionally disturbed) as do approaches which suggest modifications in educational programs based upon professional or quasi-professional developmental diagnoses. Rather, the goal is to identify observable maladaptive behaviors which appear to interfere
with learning and to arrange (rearrange?) the contingencies of reinforcement to expedite that process.

The present review will detail some of the behavior modification strategies which have utilized the teacher as the agent of reinforcement and where selected behaviors of the entire class have been pinpointed as the target behaviors. At least three such strategies have been investigated: the systematic use of teacher praise and ignoring responses contingent on "study" and disruptive behavior; the use of direct access to reinforcers such as free time and material rewards contingent on appropriate behavior; and the implementation of token reinforcement programs.

Repeated studies have demonstrated that socially desirable classroom behavior can be increased by teacher attention for appropriate behavior and ignoring disruptive acts. The rationale for the success of such a simple manipulation of contingencies has been the observation that teachers rely primarily on aversive control, especially threats of punishment, to maintain classroom order. Appropriate behavior more often than not goes unnoticed, and the teacher thus inadvertently reinforces annoying responses by his attention to them. In implementing this technique, the teacher praises and attends to good behavior or to improvement if the desired terminal response is not yet in the student's repertoire. His disruptive acts,
on the other hand, are ignored with the assumption that they will extinguish when environmental support is withdrawn and competing appropriate responses are reinforced.

Hall, et al. (1968), for example, increased the percentage of time that problem students in several elementary classrooms in a disadvantaged area spent "studying," i.e., orienting to the appropriate person or object and participating when requested, by the use of an observer in the classroom who cued the teacher to approach and praise the target students when they were studying. Madsen, Becker, and Thomas (1968) extended this finding by determining that the praise and attention aspect, rather than the ignoring one, was the important factor in the change process. In their study a Rules condition during which the teacher reminded the students of the rules was ineffective in changing the amount of time that the children spent in off-task behaviors. A Rules plus Ignoring inappropriate behavior condition actually increased disruptive behavior. It was only a Rules, Ignore, and Praise for appropriate behavior condition which led to increments in on-task behavior. The effectiveness of teacher attention in reducing disruptive behavior has been reported for secondary school classrooms (McAllister, Stachowiak, Baer, and Conderman, 1969). Taking the opposite tack, Thomas, Becker, and Armstrong (1968) produced a large increase in disruptive behavior in a normally well-behaved elementary school class
by asking the teacher to discontinue praise altogether and to attempt to control the class by disapproving statements entirely.

While praise and other social reinforcers connected with teacher behavior have been effective in producing desired classroom behavior in a number of cases, at least two problems with this strategy have been cited. First, Ward and Baker (1968) noted that this technique was not uniformly successful for all students in their sample ("normal" first-graders). The withdrawn and inattentive children seemed little affected by teacher approval and attention, perhaps because of the absence of behavior approximate to the desirable terminal response which the teacher could shape. Staats (1970) also has noted that the attentional and work behaviors of children with special learning problems, e.g., emotionally disturbed or disadvantaged learners, are often poorly maintained by social reinforcers usually effective in normal populations such as grades, being correct, and teacher approval and censure.

O'Leary, Becker, Evans, and Saudargas (1969) report another problem with the praise and ignore strategy. They observed that where peer approval is maintaining inappropriate behavior, the teacher's ignoring the disruptive response may be ineffective in extinguishing the behavior. Adolescents particularly have been found much more responsive to peer attention than to teacher or administrator approval
or censure (Meichenbaum, Bowers, and Ross, 1968). For these reasons researchers seeking to promote more efficient academic work and social behavior have sought to introduce material reinforcers into the classroom environment, either directly contingent on the desired behavior or mediated by token programs.

The use of direct access to reinforcers such as free time and material rewards is a second general strategy which has been effective in improving social and academic behavior in the classroom. Hopkins, Schutte, and Garton (1971) set up a system whereby students in two lower elementary school classes could go to a special playroom as soon as they had completed an assignment. The rate at which these students printed or wrote daily copying assignments increased reliably with no accompanying decrease in accuracy when they could enter the play area upon completion of the assignment at an individually programmed level of acceptable accuracy. Even with only intermittent checks every few days, the effects were maintained in a similar study with kindergarten subjects (Salzburg, Wheeler, Devar, and Hopkins, 1971). Other researchers have demonstrated that this procedure can readily redistribute teacher contacts between good and less able students (Sanders and Hanson, 1971). Since the better students are the first to complete their assignments and to exit for the play area, the teacher can focus more attention on
Osbourne (1971) has decreased the instances of out-of-seat behavior by awarding free time for the incompatible response of staying in one's seat. Noise levels have been reduced during study hall periods by making extra recess contingent on the appropriate behavior (Schmidt and Ulrich, 1969).

Another means of engineering more powerful reinforcers than teacher attention into the classroom, and the third general strategy for improving classroom behavior, has been the development of token reinforcement programs. Typically in these programs, tokens such as checks, poker chips, stars, or points are delivered to each student by the teacher contingent on previously agreed upon appropriate behaviors (e.g., sitting in one's seat, obtaining permission before speaking, or a certain number of math problems completed correctly). At some later time the tokens can be exchanged for backup reinforcers (i.e., objects of concrete rather than symbolic value to the subjects) such as toys, candy, and special privileges. In most cases the student is offered a choice of several reinforcers prejudged attractive to him. Additionally the teacher is instructed to praise the student as the token and backup reinforcement are delivered, as well as frequently during the day. This pairing of concrete reinforcement and teacher approval is assumed to enhance the value of the latter as a social reinforcer such that praise alone
will be sufficiently rewarding to maintain the appropriate behavior as the token system is gradually withdrawn, although this assumption apparently has not been empirically investigated.

The token economy seems to have a couple of advantages over direct reinforcement with material rewards. First, there is no distraction effect; children do not stop their work to consume tokens as they might edibles. Also the use of a variety of backup reinforcers avoids the problem of satiation on any one reinforcer, as well as controlling for daily and intersubject variation in the appeal of the reinforcers.

Numerous studies have demonstrated that a token reinforcement program is effective in modifying both social behavior and academic performance. O'Leary, et al. (1969) developed a token economy in a second-grade classroom of disadvantaged students who initially exhibited disruptive behavior (defined as wandering about the room, aggressive acts toward other children, disturbing others' property, operant noise, and talking) an average of 54 percent of the time. During the token phases the teacher rated each child on how well he followed a set of class "courtesy" rules and placed the points that he had earned in a notebook on his desk. The points were exchangeable for candy, pennants, dolls, comics, and toys. Disruptive behavior dropped to an average of 30 percent under the token arrange-
ment. Other studies have shown similar results for the effectiveness of a token reinforcement program in reducing disruptive behaviors (e.g., Bushell, Wrobel, and Michaelis, 1968; Packard, 1970; Zimmerman, Zimmerman, and Russell, 1969). Gains in academic behaviors have been reported by Clark, Lachowicz, and Wolf (1968); Ferritor, Buckholdt, Hamblin, and Smith (1972); Miller and Schneider (1970); and Wolf, Giles, and Hall (1968).

Fewer studies have dealt systematically with the effects of variations within the token economy itself. Some attention has been directed toward assessing the relative efficacy of group versus individual contingencies, since the keeping of detailed records and awarding of tokens on an individual basis is somewhat time-consuming. Herman and Tramontana (1971) reported that a group reinforcement technique was at least as effective as individual reinforcement in modifying disruptive behavior in a Head Start classroom. Other investigators (e.g., Phillips, Phillips, Fixsen, and Wolf, 1971) have suggested that an important dimension requiring clarification in the arrangement of token economies involves the relationship between the tokens and behavior.

Token reinforcement programs may be entirely "positive" with tokens awarded for appropriate behavior and no penalty for undesirable responses. On the other hand, the economy may involve only "negative" contingencies, where the
individual starts the day with a fixed number of tokens and then systematically loses them for inappropriate responses. Or the token program may be "mixed" with opportunities both to earn tokens for appropriate behaviors and to lose them for the corresponding negatively sanctioned behavior.

Behavior modifiers employing token programs in classroom settings traditionally have emphasized the "positive" model. The student is rewarded with tokens and praise for a certain level of acceptable behavior while his disruptive acts are ignored. Occasionally time-out procedures (e.g., the child is removed from the room and therefore the opportunity to earn tokens for a short time) are cited as additional control measures (e.g., Birnbrauer, et al., 1965), but quite often there is a reluctance to use a token loss contingency (Vernon, 1972).

This reluctance to use "negative" or "mixed" contingencies in token programs no doubt rests on several considerations. One result of the use of traditional types of punishment (e.g., the presentation of aversive stimuli such as physical punishment, shock, and unpleasant auditory feedback) seems to be the conditioning of fear-arousing responses and avoidance of the punishing agent and situation. Vernon (1972) noted that the more frequently punished students are more likely to drop out of school; and when in class they are reluctant to approach the teacher, even sitting as far in the back of the room as possible. In this
same vein, a great deal of literature has established that conditioned emotional responses in the form of test anxiety may well be inhibiting to learning processes (e.g., Sarason, Davidson, Lighthall, Waite, and Ruebush, 1960). Second, the punishing effect of low grades and teacher disapproval received by lower ability students who may be performing as well as they are able probably suppresses their academic responses. Moreover, while punishment establishes a link between the punished response and aversive consequences in the environment, it gives no cue to the appropriate response.

In contrast to the presentation of aversive stimuli, the withdrawal of positive reinforcement as a consequence of negatively sanctioned behavior seems to suppress those responses without the attendant emotional responses. In fact, Bandura (1969) notes that the withdrawal of positive reinforcers tends to maintain, and may even increase, approach tendencies toward the change agents. He further suggests that the use of withdrawal of positive reinforcers for undesired behavior adjunctive to positive reinforcement for appropriate behavior (the "mixed" model) may be superior to simply ignoring the negatively sanctioned behavior (the "positive" model), if the inappropriate response is being maintained by rewards dispensed by someone other than the change agent. In the mixed model, for example, the personal penalty of losing tokens and therefore some kind of material
reward or privilege would be pitted against the reinforcing effects of peer approval.

Laboratory studies have shown that response-cost contingencies where tokens exchangeable for money are removed for disapproved behavior suppress that behavior (e.g., Weiner, 1962). Similar suppressive results have been reported in token programs in applied settings (e.g., Phillips, 1968). Only two studies, however, have compared a system of response-costs and positive reinforcement with one of positive reinforcement alone within the same group of subjects; and neither of these studies was conducted in a classroom.

Boren and Colman (1970) observed a deterioration in the desired terminal behavior when a mixed contingency system was substituted for a positive system. These researchers were interested in increasing the resident attendance at an early morning unit meeting in a psychiatric ward for delinquent soldiers. Even though points were given for attendance (these tokens could be exchanged for snack bar, pool, and T.V. privileges as well as weekend passes), at times 30 percent of the men failed to attend. A new contingency was imposed calling for a fine to be levied for nonattendance in addition to positive reinforcement for attendance. On the first day after the announcement of the fine system, attendance dropped to a new low of 50 percent. On the second day no men attended;
there were whispered threats of "rebellion" and several men went AWOL. After a week of increasing disruption, the fines were rescinded and attendance increased to its original level. Boren and Colman noted that these men had been assigned to the ward for aggressive acts directed at officers and for refusal to follow orders, and they speculated that mixed contingencies might well be effective in other settings and with other populations.

In contrast to Boren and Colman's (1970) results, Phillips, et al. (1971), comparing mixed, positive, and negative token-behavior models, found that the mixed model was the most efficient and the positive model the least efficient in modifying some aspects of adolescents' interest in current events. In a community-based home for delinquent boys, they operationally defined interest in current events as the subjects' watching the Huntley-Brinkley newscast each evening and performance on a 7-to-9 item news quiz taken from that night's broadcast. Points could be earned or lost contingent on the number of quiz items correct or incorrect in a series of complicated manipulations. The data indicated that only one condition, the response-cost plus positive arrangement, resulted in all the subjects watching the newscast each evening. The negative contingency condition was less efficient than the mixed, but more effective than any of the several variants of positive only contingencies. Interestingly, regardless of the total
number of points awardable for each item in the positive only conditions, the number of boys who watched the news each evening dropped to zero in a couple of days. The effectiveness of the various contingencies on comprehension of the broadcast was somewhat more equivocal. The percentage of correct responses on the quiz among those who watched the news did not change appreciably across the conditions.

In the same study a mixed contingency model was effective in maintaining an almost perfect level of cleanliness in the subjects' rooms. Phillips, et al. (1971) speculated that while a positive contingency arrangement may be more effective in increasing appropriate behaviors that have an initially low baseline rate, and a negative contingency most efficient in eliminating inappropriate behaviors with a high baseline rate, a mixed arrangement may be most effective in maintaining behavior at a certain desired level.

The control of disruptive behavior in the classroom seems to be a problem of maintaining appropriate behavior at a high level rather than shaping new responses since most subjects have "study" behavior present in their repertoire although it may be emitted only sporadically. Based on Phillips, et al.'s (1971) conclusion regarding the relative effectiveness of positive reinforcement alone and positive reinforcement plus response cost contingencies,
further systematic investigation of the two techniques in a classroom token reinforcement program seems warranted. Moreover, while behavior modifiers have deemphasized the notion of diagnostic classifications, it may be that a mixed contingency token system is especially effective with retarded subjects. Vernon (1972) suggests that the use of a token economy can create sensitivity to the existence of behavior-reinforcement contingencies. He notes that the average adult and most children are aware of approval or disapproval of their actions and adjust their responses accordingly. This ability, however, seems to be lacking to some extent in most mentally retarded children. They simply do not seem to use the cues which society offers as guides for their behavior in various situations (Leland and Smith, 1965). Vernon (1972) further contends that the consistency and the immediacy of behavior-reinforcement relationships in the token reinforcement program can develop an awareness of social stimuli so that the child learns the general strategy of modifying his behavior with respect to the immediate environmental consequences. A mixed contingency token arrangement with its salient cues for inappropriate responses as well as for socially sanctioned ones, therefore, would seem to be indicated for dealing with mentally retarded children in various social situations, and especially mentally retarded populations in classroom settings.
This notion, interestingly, is largely untested. Broden, Hall, Dunlap, and Clark (1970) demonstrated the use of response cost procedures in a special education classroom; but no studies have systematically compared mixed and positive token arrangements in this setting. More precise delineation of token contingencies which are effective in modifying the classroom social behavior of retardates seems important in view of the current disillusionment with traditional special education programs.

Placement in self-contained classes has been the usual educational arrangement for mildly retarded children and adolescents in American society. The American Association on Mental Deficiency (AAMD) defines mental retardation as "subaverage general intellectual functioning which originates during the developmental period and is associated with impairment of adaptive behavior" (Heber, 1961, p. 3). In practice, however, only the first aspect, intelligence as measured on a standardized test, has been used as the criterion for placement in a special class.

These classes for the educable mentally retarded (EMR) ordinarily include children with measured intelligence in the range from 50 to 70 and have operated under the assumption that retarded children have some type of learning deficit which distinguishes them from normal students (e.g., some of the hypothesized differences are an inability to
think abstractly and a slower rate of learning). Often these assumptions have led educators and school systems to introduce significant modifications in the regular educational program for EMR classes. Generally, these changes in program have included at least one of the following: (1) Providing the normal curriculum at a reduced rate, (2) Seeking curricula designed exclusively for special class use, (3) Establishing professional training which involves extensive study of the characteristics and etiologies of exceptional children, and (4) Restructuring special classrooms by either reducing the pupil-teacher ratio or by eliminating much of the usual visual and auditory stimulation (Haring and Lovitt, 1967, p. 33).

Surprisingly, early enthusiasm about the potential of special education classes for improving the academic achievement of EMR children has not been supported very well by subsequent experimental verification of the efficacy of such programs. Many of these studies, unfortunately, have suffered from methodological and sampling problems which would bias against superior performance in special classes (cf. reviews by Guskin and Spicker, 1968; Kirk, 1964; and Quay, 1963). Evidence from other studies seems to indicate that the academic performance of retarded students placed in special education classes is, at best, only comparable to that of their counterparts who are retained in regular classes; and in a number of studies, special
education students do much poorer.

Furthermore, the common notion that the personal and social adjustment of retarded children is better in special classes than in regular classes has little empirical support (Gardner, 1966). There is, in fact, increasing concern that special class placement and its attendant stigma may have a devastating effect on the retarded individual in terms of its impact on his self-esteem and his teacher's expectations regarding his abilities.

There have been frequent reports of the negative stereotypes associated with the label "retarded" by persons of normal intelligence (e.g., Guskin, 1963; Wilson, 1970), but little research has been focused on the attitudes of the retarded themselves regarding that label. Edgerton (1967), in a follow-up of previously institutionalized mentally retarded adults, found that his subjects were reluctant to discuss their former placement or denied it altogether. His sample preferred nonretarded friends and had few ties with former acquaintances who also had been released from the institution. Virtually all believed that their commitment had been an error.

This tendency to reject the label "retarded" has been noted among students in special classes also. Jones (1970) reported that for a small sample of high school males from an EMR class, 17 of 23 denied that they were members of a special education class on a questionnaire about their
school work. When later asked why they had falsified the information, most stated that it was in an effort to avoid ridicule. Special education classes were disliked because members were often made fun of. These boys reported that when regular class students queried them about the procedures in the special class, they were always careful to say that the work and grading system were the same as in regular classrooms.

The stigmatizing effects of a "retarded" label are also important because of the hypothesized effects of teachers' expectations about students' abilities on their ultimate levels of achievement. Rosenthal and Jacobson (1968) rather dramatically focused attention on the problem of self-fulfilling prophecy in the classroom. In their study, teachers in an elementary school were given a list of children from whom to expect a rather sudden spurt of intellectual growth. These "late bloomers" were, in reality, randomly assigned to that condition. In follow-up testing of the sample at the end of the school year, it was found that mean IQ gains for the experimental subjects (i.e., the "bloomers") were significantly higher than those for the non-labeled control group, at least in the first and second grades.

Rosenthal and Jacobson's (1968) results have been challenged on a number of methodological grounds (cf. Barber and Silver, 1968; Jensen, 1969; and Thorndike,
follow-up studies in large classrooms have failed to replicate their results (e.g., Clairborn, 1969; Fleming and Anttonen, 1971; Gozali and Meyen, 1970; and Jose and Cody, 1971). Follow-up studies in micro-teaching situations, on the other hand, have yielded some results which support the hypothesis of teacher expectancy effects; and researchers have begun to examine how such expectancies might be mediated. Beez (1968) asked teachers (graduate students in Education) to teach children a series of symbols in a one-to-one situation. Half of the teachers were given a psychological evaluation of the child which indicated a good prognosis for learning, while the other half were led to believe that their child was a poor learner. The results indicated that among those children labeled as having a good intellectual potential, 77 percent learned five or more symbols, whereas only 13 percent of those reported to have poor ability performed at that level. Moreover, teachers given positive feedback about their students attempted to teach more items than did teachers whose students supposedly had a poor prognosis for learning.

Rubovits and Maehr (1971), also studying the effects of teacher expectancies, found that equal amounts of attention were given "gifted" and "nongifted" sixth and seventh graders by student teachers in a micro-teaching situation. However, the pattern of the interactions was different; students
labeled "gifted" were given more praise for their responses and were more often requested to participate.

These studies demonstrating differences in teacher behavior as a function of expectation were carried out in rather artificial situations. The "teachers" were either undergraduate or graduate Education majors who were given little other information about the situation than what the experimenter wished and were allowed only one teaching encounter. It may not valid, therefore, to generalize the effects to a classroom especially since these students probably were anxious to please the experimenter. Still the findings rather provocatively indicate that, at least in some situations, biases in teacher expectancy do affect subsequent teacher-student interaction as well as the pupil's performance. Moreover, Beez' (1968) finding that teachers of children with "poor" learning prognoses presented educational materials at a slower rate than did teachers of high potential children might help to explain the fact that EMR children retained in regular classes general do better on standardized achievement tests than retarded children placed in special classes (Guskin and Spicker, 1968). Apparently these children are more likely to be exposed to larger quantities of material.

Besides the disappointing results in academic performance for EMR students in special classes and the potentially stigmatizing effects of such placement, some insightful
educators also are concerned that special class placement is far too often an exclusionary procedure rather than the initiation of a remedial process (Gallagher, 1972). Although many special educators have seen special class placement as a temporary arrangement until certain academic and behavioral problems are solved, recent research has indicated that in a number of large city school systems, less than ten percent of the children placed in these classes are ever returned to the regular curriculum (Gallagher, 1972). Teachers and administrators alike seem only too eager to relegate their most bothersome students to a self-contained special class.

The presence of behavior problems in addition to academic retardation may well be the factor which points a child toward special class placement. Several studies have shown that there are a number of mildly retarded individuals in regular classes who have failed to come to the attention of the school authorities (e.g., Mercer, 1971). The EMR students who are ultimately placed in special programs exhibit more behavior problems than their retarded peers who remain in regular classes (Kirk, 1964). This observation suggests that the visibility of the retarded child in the normal classroom in terms of his social behavior might be an important factor in the teacher's decision about whom he will refer for consideration for special classes.
It is precisely such disruptive behaviors which have been especially amenable to modification by the use of a token reinforcement program. Furthermore, these inappropriate behaviors (e.g., the inability to work independently, aggressive acts toward other children or the teacher) are similar to current notions of adaptive behavior. Adaptive behavior has been defined as the degree to which the individual successfully copes with his environment in terms of his ability to function independently and to meet culturally imposed standards of personal and social behavior (Heber, 1961; Leland, Shellhaas, Nihira, and Foster, 1967). Therefore, it seems that impairment in adaptive behavior, while not a de facto criterion for being placed in special classes in public schools, operates subtly in the selection process.

Although token reinforcement programs have been shown to be effective in reducing baseline levels of maladaptive classroom behavior of both normal and retarded children, researchers have not demonstrated whether such programs can improve and maintain social behavior among EMR students in a special class at a level comparable to their normal peers in regular classes. Such a comparison seems warranted if token reinforcement systems are to play a remedial role in mildly retarded students' regaining access to regular classes.
Aims of the Present Research

Following from the findings of Phillips, et al. (1971) and the speculations of Vernon (1972), this study sought to examine the relative effectiveness of two contingency arrangements, positive reinforcement only and positive reinforcement with response costs, in a classroom token reinforcement program in modifying the maladaptive behavior of educable mentally retarded adolescents in two special education classrooms. Second, it sought to assess the usefulness of a token reinforcement program as a remedial tool in general by examining its effectiveness in developing and maintaining a level of appropriate adaptive behavior comparable to that expected from normal students in a classroom in the same school.
Chapter II

METHOD

Subjects

Fifteen members of two special education classes in a junior high school, four girls and eleven boys, served as subjects. Half were selected from a group which their teachers designated as the most disruptive and/or having the poorest study habits in the class; and half were randomly selected, all with the additional requirement of better than average attendance records. The teachers were not informed about which students were included in the sample. Ss ranged in age from 12 years, 6 months, to 16 years. Their mean measured intelligence (WISC Full Scale) was 68.2 with a range from 48 to 77, and all had been in a class for educable mentally retarded students for at least five years.

A wide range of maladaptive behaviors was described by the Ss' present teachers and their cumulative folders. Four of the male Ss had frequently been sent to the office for open defiance of the teachers and fighting with other students and had been suspended several times as a consequence. The other students were not cited for similar instances of aggressive behavior, but nevertheless exhibited responses which annoyed both the teacher and the rest of the class and interfered with the learning process. They were
variously described as "anxious," "immature," "hyperactive," and "disorganized," as well as having "short attention spans" and "irrelevant speech."

Eight students, four girls and four boys, were similarly selected from a so-called "normal" English class in the same school to serve as a comparison group. Half of these adolescents were described as often disruptive and failing for the term. Of the four randomly selected, one was a hall patrol who supervised movement between classes, one was a spearhead of the Chicano group in the school, and one served as the class president.

Setting

The token reinforcement program was developed in two special education classes, each with 17 to 18 students, during the second semester of the academic year in a junior high school in Southern California. The classes met for the first four periods of a six-period day; the first period regularly involved social studies and health units, the second reading and language arts lessons, the third math, and the fourth covered typing and miscellaneous science projects, films, current events discussion, etc. The students were grouped heterogeneously with regard to intellectual ability during the first and fourth periods, while for the second and third periods the students were divided into upper and lower tracks on the basis of academic skills. Most of the subjects made class changes during the day, but some
remained in the same teacher's classroom because they were considered "unable to handle" room changes.

Although the teachers of the two classes did not collaborate in planning academic programs for the students, the daily procedure in both classes was similar. A general assignment was given to the entire class at the beginning of each period (e.g., to continue working in reading or math workbooks) with each child then expected to work independently at his own rate. The materials were presented in a traditional, non-programmed format; the teacher gave help as needed in class, collected the books or papers at the end of the period, and then corrected the work for the following day.

Just before the study was begun, one of the teachers initiated a cooking unit for the students who were best behaved and who had good attendance records. During the periods devoted to cooking, the pupils who failed to meet the requirements to cook were sent to the other classroom for a study hall during which they were expected to make up work or to begin the day's assignment. Generally no teacher assistance was given during these periods.

The teachers of both EMR classes had teaching credentials for special education, but each had a somewhat different orientation to the classroom. Ms. A., in her third year of teaching, expressed enthusiasm about her job and often sought outside activities which involved her students,
as taking the girls to the beauty shop and inviting the class to her home for parties. During the Baseline period she was observed to maintain classroom control by joking warnings about minor disturbances; for example, to a loud groan, she might respond, "Are you sick? Do you need to see the school nurse?" To continued disturbances she generally asked the student to put away his work and to report after school to complete it. In cases of physical or verbal aggression toward herself or another student, she immediately sent all parties to the office to have the matter straightened out. The classroom atmosphere was fairly relaxed; the students were allowed to chew gum, they could move about the room to get supplies or to throw away paper, and occasionally there were times when they were given permission to talk quietly at the end of the period when they had completed their assignments.

Ms. B., in her second year of teaching, often commented that if she had known teaching would be such a battle of wills, she would have chosen another profession. She had no contact with her students outside of class and maintained rigid rules during the day. No student was allowed to leave his chair or to talk without specific permission. Warnings were issued at the first incident of noise, and then the student in error was given detention. During the Baseline period at least one third of the class was assigned detention for various offenses each day. Four were blocked off from
the view of the class by plywood partitions reportedly because they provoked the rest of the students.

The comparison "normal" class was randomly selected from several that the principal had suggested as representative of the school. This ninth-grade English class was run similarly to the EMR classes in that there was little formal instruction to the group as a whole. The course used a programmed text, and the students were expected to work independently, asking the teacher for help as they needed it. The one class rule that the students not disturb others who were working was generally enforced by verbal reprimands.

The junior high school in which the study was conducted served a racially mixed neighborhood. Following racial trouble several years before the study, a new principal had instituted an elaborate system of discipline. For a first minor offense (e.g., tardiness, failure to complete work, bothering others by talking), the student was given a warning by the teacher, followed by a letter to the parents of the student if the behavior continued. For major offenses (e.g., fights, vandalism), the student was sent to the office for counseling, then physical punishment or suspension. Twenty days of suspension put the student on an "automatic suspension" list which gave the teacher permission to send him directly home from class without the usual office mediation.
Observation Technique

Observations of the EMR subjects were made three periods a day, five days a week throughout the study. Ms. B's class was observed during the first period each day since she oversaw the study hall when part of the students went to the kitchen to cook; during the second and third periods each day Ms. A's classes were observed.

Each class period was divided into six equal time segments, varying from eight to ten minutes with the length of the period. During each segment the occurrence of specific disruptive or appropriate behaviors described below was scored for one subject on a ten-second interval time-sampling basis. Each minute was divided into six ten-second intervals, and the observer noted the occurrence of appropriate or inappropriate responses during that interval. The order in which the subjects were observed varied randomly within and over periods, but an effort was made to observe each child two periods a day. During the study halls the children who were part of the sample were systematically observed.

The normal subjects were similarly observed, i.e., on a ten-second interval time-sampling basis, during the third period when their class regularly met for three weeks during the last month of school.

The experimenter served as one observer five days a week throughout the study. A senior in Psychology at a
nearby college who was paid for her help served as a second observer three days a week. She had some knowledge of behavior modification techniques and, of course, could not be blind to the Baseline and Token conditions but was never present during explanations of the various token arrangements. The observers were represented to the classes as prospective teachers who would be taking some notes on how the teachers instructed their students as part of a term project. Although observers are usually cautioned to minimize responses to the subjects, it was felt that a more natural procedure was to show affect when the whole class responded (e.g., to something amusing) but to ignore individual attention-getting responses. The observers sat at the side of the room where the entire class could easily be seen; they attempted to avoid eye contact if it seemed that the subject was aware of being observed. During the breaks between classes, students occasionally asked the observers questions; these inquiries were answered pleasantly but as perfunctorily as possible.

Prior to the Baseline period the observers spent a week in the classrooms to accustom themselves to the observation technique and the students to their presence. During the first week of the study, reliability checks were made by concurrent observation of the same subjects over all three periods; thereafter, reliability checks were made for one period only each day with the order randomly varied.
Behavior Definitions

The classes of behavior selected for observation were adapted from O'Leary and Becker (1967) who categorized behavior occurring with some frequency in the repertoires of problem children. Noted also by the teachers in the present study as responses which interfered with good classroom learning conditions, six categories of inappropriate behaviors were defined as follows:

**Out of seat (O):** The student breaks all contact with his seat without permission. No part of the body is touching the chair or table top.

**Talking (T):** The student makes a verbal statement without permission. This may be speaking out in class without raising his hand or talking to his neighbors, if the vocalization is audible to the observer. Reading aloud or talking to oneself while studying is not scored. (Occasionally Ms. A. allowed quiet talking at the end of the period; if announced as permissible, this behavior was not marked as an occurrence.)

**Making noise (N):** The student makes some type of operant sound other than a verbalization which is audible to the observer. This includes tapping pencils, beating on the desk top, stamping feet, or crumpling paper repeatedly as when cleaning out one's desk. Also included are instances of vocalizations other than verbalizations (e.g., operant coughing, humming, groans). Noise from dropping objects or
moving one's desk is not scored as an occurrence unless the student is throwing something at someone or attempting to get his attention.

**Disturbing others (D):** The student touches another student (directly or indirectly) or an article that the other student is holding or has on his desk. Also included is gesturing or posturing in the direction of another student in an attempt to get his attention. This category excludes talking to another student.

**Orienting response (O):** The student orients his eyes and head in the direction of the window or open door (excluding a brief glance, as at a noise). The student turns to the person in back of him when the teacher is standing at the front of the room.

**Time off task (X):** The student is engaged in some task other than the assigned one. Examples are staring into space for longer than thirty seconds, failing to follow instructions within one minute after they are given, doing math homework during reading class, or cleaning out one's desk without permission. Instances of failing to look at the teacher as he instructs the class are not scored in this category if other behaviors of the student (e.g., asking or answering questions) indicate that he is attending.

**Appropriate behavior (A):** was defined as on-task responses, i.e., answering questions of the teacher, raising one's hand and waiting for the teacher to respond, writing
answers to or reading material which was assigned. Appropriate behavior had to occur for the full ten-second interval to be scored as such.

The observers recorded each category of behavior which occurred during the interval regardless of how many other categories had already been noted for that interval; Out of Seat, Talking, Noise, Disturbing others, and Orienting were compatible with each other. No class of behavior was scored more than once during an interval, since it was found that Talking and Noise were not easily quantified as frequencies.

**Experimental Conditions**

Since the order of presentation of the token conditions could not be counterbalanced across the EMR classes because they were not self-contained, the five phases of the study were as follows: Baseline period, Token+ condition, Withdrawal of the token program, Token* condition, and Token+2. In the comparison normal class, Baseline data only were gathered.

**Baseline.** Following one week during which the observers increased their recording skills and the students became accustomed to the observers' presence, the Ss were observed under normal classroom conditions for 13 days over a three-week period. The teachers were instructed to conduct class as usual. During the last week of the Baseline period in
the special education classes, the experimenter met several times with the teachers to discuss behavioral principles and to outline how the token reinforcement program could be most conveniently implemented in their classrooms.

Token+1 condition. The positive reinforcement-response cost contingency token program was the first condition because Ms. B was not at all confident that a system of positive reinforcement only would continue to control the class as well as her usual method.

Each teacher explained the token reinforcement program to her class, loosely following a script prepared by the experimenter as follows:

"Ms. A(B) and I have been talking during the last few days and are considering a different way to run the class. Our job right now is to be your teachers, and we get a paycheck for doing it--coming to school every day, correcting papers, making out your report cards.

You have a kind of job, too. Right now you're expected to come to school, do your work, and follow the class rules. You get paid in certain ways for your work, too--your grades, being passed to a new class. But we have been thinking that it might be a good idea to set the class up more like a regular job. So here's what we would like to do if you agree.

Each period you will be able to earn points depending on how well you follow the class and school rules, and then you can exchange these points for certain things every few days.

Here's how the program will work. You each will have a card with your name on it. At the end of each period, either I or Ms. A(B) will stamp it with the number of points you've earned that hour. You can earn points according to how well you follow these class rules."
A sheet with each of the rules and the points possible was given to each student. (See Appendix A for an example.) After explaining the rules for which points could be earned, the teacher continued:

This means that if you follow all the rules, you will earn 10 points. Since there are four periods a day, you can receive up to 40 points every day.

But you can also lose points if you fail to follow the rules. You can lose points for these things. (See Appendix A for examples.)

Let me show you how this could work. If, for example, you were talking to your neighbor, but you followed all the other rules, you would receive 8 points for staying in your seat, raising your hand before talking, leaving other people alone, and good hall behavior. However, since you talked to your neighbor during that class period, I would have to remove 2 of those points that you had earned, so that you would receive only 6 points for that period.

Other examples were given, and the teacher had the students themselves figure out the number of points earned given a particular situation. The teachers further explained that they would keep a record of each student's behavior and that the points would be awarded just before the bell rang to end each period.

The teachers were instructed to use liberal praise for appropriate behavior during the day. For the first week of the roken program each student was given a quiet warning when he misbehaved. Following a second offense during the same period, the teacher went to his desk and quietly told him that he would lose points, but the actual removal did
not take place until the end of the period. At that time the teachers were instructed to compliment the children on the points that they had earned and to explain briefly why they lost points, if any. When a student lost points, the teachers were to cross out previously earned points with a red pen rather than simply subtract the points lost from the total earned for that period.

The points were exchangeable for materials which the teachers could order through the school store—composition books, colored pencils and chalk, water colors, erasers, rulers, clay, bookmarkers, calendars, piggy banks, and beads—as well as edibles supplied by the experimenter. About once a week a new backup reinforcer was introduced; other items used were decks of cards, rock and sports magazines, and certificates for special lunches which the students cooked and ate themselves in their kitchen (many of the subjects did not like to eat in the cafeteria area with other students in the school). An attempt was made to keep the backup reinforcers in each of the rooms comparable, but Ms. A often brought extra treats for her children.

During the initial four days, the children were eligible for the reinforcers at the end of fourth period each day. Thereafter, points were exchanged for backup reinforcers twice a week. For this condition there were 17 observations over a four-week period.
Withdrawal condition. For this phase of the study, the children were told that the school system was running low on funds, and the token system would have to end since there was no more money for prizes. Each teacher emphasized that she was pleased about how well the class had been working in the last few weeks and that she expected them to continue to do as well. The students were allowed to spend their remaining points, and then the cards and prize tables were removed from the room. This phase lasted for one week during which there were observations on five days.

Token+ condition. For the positive reinforcement token arrangement condition of the study, the students were informed that some money had finally been found and that the token program could be started again, but that the rules would be slightly changed. There would no longer be any fines; in fact, a student who earned all ten points for the period would receive a bonus of 5 points so that he could earn 15 in all, or a possible 60 for the day. Because classwork had slumped during the Withdrawal phase, the class rules were changed somewhat so that points could be earned for the behaviors shown in Appendix B.

The teachers were again instructed to praise the students for improvement and to attend to appropriate behavior during class. Points were exchanged for backup reinforcers once a week, and the prices were adjusted upward slightly so that more points were necessary to obtain a reward.
Observations were made on twelve days over a four-week period.

**Token+2 condition.** During the final phase of the study the positive reinforcement and response cost contingencies were reinstated to note whether the effect for this condition was reliable and whether differences between the two token conditions might be due to differences in the number of points possible, since the Ss expressed a preference for the condition in which they could earn more points although the buying value was virtually the same in both.

The students were told that again penalties would be levied for failing to follow the rules. The procedure was the same as the Token+1 condition except that there were 15 points possible for each period and a fine of 5 points was given for inappropriate responses. Five observations were made over a two-week period.
Chapter III

RESULTS

Reliability

The reliability of the observations was checked by calculating the inter-observer agreement on simultaneous but independent observation records. The two records were compared interval by interval, and an agreement was scored if the same behavior category, or categories, was recorded in the same interval by each observer. Reliabilities were calculated by dividing the number of intervals in which there was agreement by the total number of agreements plus disagreements.

During the first week of the Baseline condition, reliability checks were made by concurrent observations of the same subjects over all three periods in the EMR classes; thereafter, reliability checks were made for one period only each day on which there were two observers. Inter-observer agreement in the special education classes ranged from 67 to 100 percent during the Baseline period with a mean of 94 percent. During the token reinforcement program, agreement ranged from 81 to 100 percent with a mean of 94. In the normal class one reliability check was made with agreement ranging from 78 to 97 percent for each child and a mean of 89 percent.
Classroom Behavior of EMR Students with Token Reinforcement

The dependent measure reported was the percentage of intervals in which a particular category of behavior was observed. Percentages, rather than frequencies, of intervals were reported since the length of the observation period for each child varied with the length of the period and the number of children present that day. Because there was some variance across classes, the data are presented for Period 1 (Ms. B's class), study hall, and Period 2 (Ms. A's class) settings. Data for Period 3 was omitted since the results were highly similar to Period 2 with Ms. A the teacher of both classes and most of the subjects in both.

The occurrence of Appropriate behavior among the subjects drawn from Ms. B's class is shown in Figure 1. During the Baseline period the students behaved appropriately an average of 83.9 percent of the time. Following the institution of the Token+1 condition, this average increased to 94.3 percent, an increment of about 10 percent, and the daily variability in Appropriate behavior was somewhat reduced. Upon withdrawal of the token program, Appropriate behavior dropped to a mean of 79.8 percent, slightly lower than the Baseline mean. Study behavior again increased with the reinstatement of the token system, but the mean Appropriate behavior observed under both the Token+ and the Token+2 conditions, 93.3 and 95.3 percent respectively, did not differ.
Figure 1. Mean percent of intervals of Appropriate behavior for the subjects in Ms. B's class under each treatment condition.
Graphs of the percentage of Appropriate behavior for individual students in Ms. B's class are shown in Figures 2, 3, and 4. Ss 1, 2, and 3 were selected from those designated as behavior problems by Ms. B, while Ss 4, 5, and 6 were picked randomly. The same general pattern found for the group data was observed in the individual graphs of Ss 1, 2, 4, and 6. S3 was absent so often that it is difficult to assess the program's effectiveness on increasing his Appropriate behavior; however, some indication is given by the observation that on Days 16, 24, and 25 when his on-task responses were much less frequent, he failed to pick up his card upon entering the class and Ms. B did not award him any points. Also, his Appropriate responses decreased in frequency after Day 30 when he was removed from the token program for the remainder of the year for grabbing a handful of reinforcers and walking out of class. S5's Appropriate behavior, while somewhat improved under the token system, showed more variability than the other Ss, i.e., the percentage of intervals in which on-task responses occurred dropped below 90 percent more often.

Changes in individual categories of inappropriate behavior over the experimental conditions for Ms. B's class are shown in Table 1. This class initially was so well controlled that there was little room for improvement during the token program. Only two instances of Out-of-Seat behavior, for example, were observed during the three months
Figure 2. Mean percent of intervals of Appropriate behavior for S1 and S2 in Ms. B's class under each treatment condition.
Figure 3. Mean percent of intervals of Appropriate behavior for S3 and S4 in Ms. B's class under each treatment condition.
Figure 4. Mean percent of intervals of Appropriate behavior for S5 and S6 in Ms. B's class under each treatment condition.
Table 1
Average Percentages for Specific Behavior Classes
For Each Experimental Phase in Ms. B's Class

<table>
<thead>
<tr>
<th>Behavior Classes</th>
<th>Baseline</th>
<th>Token+1</th>
<th>Withdrawal</th>
<th>Token+2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate</td>
<td>83.9</td>
<td>94.3</td>
<td>79.8</td>
<td>93.3</td>
</tr>
<tr>
<td>Time off Task</td>
<td>12.2</td>
<td>3.9</td>
<td>3.0</td>
<td>4.4</td>
</tr>
<tr>
<td>Out of Seat</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Orienting</td>
<td>1.6</td>
<td>0.9</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Noise</td>
<td>2.0</td>
<td>0.4</td>
<td>1.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Disturbing Others</td>
<td>0.4</td>
<td>0.1</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Talking</td>
<td>2.3</td>
<td>0.5</td>
<td>20.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>
of the study. The introduction of the token program, however, did seem to affect Time-off-Task and Talking behaviors. Time off Task was reduced from a mean of 12.2 percent of the observed intervals during the Baseline condition to a mean of 4 percent which remained essentially unchanged across both Token and the Withdrawal conditions. This effect was due to the fact that students no longer slept during class or simply put their heads on their desks and failed to participate. Inappropriate Talking responses remained at the initial low level of occurrence across Token conditions but increased to 20 percent when the token programs were withdrawn.

In summary, for Ms. B's class, the introduction of the token program increased Appropriate responses for five of the six target subjects, primarily because Time off Task rather than the more disruptive responses was reduced. No difference was found between the two Token conditions for any of the dependent measures.

Appropriate behavior among students in the study hall setting is shown in Figure 5. Task-relevant behavior was observed only an average of 12.3 percent of the time during the Baseline period. Although instructed to bring some work to the classroom, most students sat and played with pencils, money, or other objects or slept the entire hour. When the token economy was introduced, the students were told that they could now earn points for doing some work during the
Figure 5. Mean percent of intervals of Appropriate behavior for the students in the study hall under each treatment condition.
study hall period. The mean time spent studying rose to 80.4 percent under the Token+ condition, an increase of 68.1. This average jumped another 10 percent by the introduction of the Token* condition, but then remained about the same when the Token++ condition was reinstated (Token+ $\bar{X} = 91.0$ versus Token++ $\bar{X} = 92.6$). As in the Period 1 class, time spent studying increased as a function of the token program, but there was no consistent difference in Appropriate behavior between the two Token conditions.

Surprisingly, withdrawal of the token program did not result in a concomitant decrease in Appropriate behavior ($\bar{X} = 97.5$). However, only one study hall class was held during the Withdrawal period, and the students who happened to be present that day had better than average study performance.

Individual graphs of Appropriate behavior are not shown for the study hall setting since the students who composed the group varied from day to day (inclusion in study hall rather than the cooking class was based upon one's attendance the previous week of class). Except for a decrease in Time-off-Task behavior as a function of the token economy, inappropriate behaviors were little changed over the conditions (see Table 2 for the average percentages of specific behavior classes). During the Baseline period sitting or sleeping comprised an average of 67.5 percent of the intervals observed. Time off Task dropped to a mean of
Table 2

Average Percentages for Specific Behavior Classes
For Each Experimental Phase in Study Hall

<table>
<thead>
<tr>
<th>Behavior Classes</th>
<th>Baseline</th>
<th>Token+&lt;sub&gt;1&lt;/sub&gt;</th>
<th>Withdrawal</th>
<th>Token+</th>
<th>Token+&lt;sub&gt;2&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate</td>
<td>12.3</td>
<td>80.4</td>
<td>97.5</td>
<td>91.0</td>
<td>92.6</td>
</tr>
<tr>
<td>Time off Task</td>
<td>67.5</td>
<td>17.9</td>
<td>0.0</td>
<td>7.0</td>
<td>7.6</td>
</tr>
<tr>
<td>Out of Seat</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Orienting</td>
<td>3.3</td>
<td>1.3</td>
<td>0.5</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Noise</td>
<td>5.3</td>
<td>0.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Disturbing Others</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Talking</td>
<td>1.0</td>
<td>0.1</td>
<td>3.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
17.9 percent under the Token+ condition and further decreased to 7.0 and 7.6 percent with the Token+ and the second Token+ conditions respectively. Noise and inappropriate Orienting behaviors likewise decreased during the Token conditions; but since the baseline rates were so low, the drops were minimal.

Focusing now on Ms. A's second period class, Appropriate behavior data for the Ss as a group are shown in Figure 6. On-task behavior occurred an average of 75.4 percent of the time during the Baseline condition; on only one of the observation days did Appropriate behavior for the group rise above 90 percent. During the Token+ condition, the mean on-task behavior increased to 92.3 percent, an increment of 16.9 over the Baseline mean; an average per day of 90 percent or above was observed on 14 of the 17 days of the condition. For the two days on which there was a substitute teacher and the token program was not in effect, Appropriate behavior dropped to 48 and 43 percent. Withdrawal of the token program, contrary to the result in Ms. B's class, caused only a 4.3 percent drop to an average of 88.0 percent on-task behavior. As in both the Period 1 and study hall classes, differences between the token systems were not apparent (Token+1 \( \bar{X} = 92.3 \), Token+ \( \bar{X} = 94.8 \), and Token+2 \( \bar{X} = 89.8 \)).

Individual graphs of the percentage of Appropriate behavior per observation period for the 12 Ss drawn from
Figure 6. Mean percent of intervals of Appropriate behavior for the subjects in Ms. A's class under each treatment condition.

© Substitute Teacher
Ms. A's class are shown in Figures 7, 8, 9, 10, 11, and 12; these data are a composite of the behaviors observed in Periods 2 and 3 since each subject was randomly observed in one of the two classes each day and responses were similar in both. Three Ss were observed both in Ms. A's and Ms. B's classes—Ss 3, 4, and 5; and they appear identified similarly in the individual graphs for students in each of the classes. Ss 5, 13, 14, and 15 were selected from the group Ms. A designated as most disruptive.

The data for nine of the twelve Ss closely approximated the results from the group means discussed earlier; the token system increased the occurrence of Appropriate behavior with no difference between the two conditions. It is interesting that S3 who was removed from the token program after Day 48 did not show the same drop in Appropriate behavior that was evident in Ms. B's class. Three students in the sample—Ss 5, 14, and 15—did not appear as responsive to the token reinforcement program; their on-task behavior fell below 90 percent at least twice as often as the other students' behavior. S5 and S15 sat next to each other and often collaborated on their work. At the teacher's insistence, S15 was removed from the token program after Day 44, following a series of incidents culminating in his feigning illness to leave school one day. S14, unlike the rest of the class, was given individual assignments since his achievement level lagged several years
Figure 7. Mean percent of intervals of Appropriate behavior for S3 and S4 in Ms. A's class under each treatment condition.

○ Substitute Teacher
Figure 8. Mean percent of intervals of Appropriate behavior for S5 and S7 in Ms. A's class under each treatment condition.

© Substitute Teacher
Figure 9. Mean percent of intervals of Appropriate behavior for S8 and S9 in Ms. A's class under each treatment condition.

© Substitute Teacher
Figure 10. Mean percent of intervals of Appropriate behavior for S10 and S11 in Ms. A's class under each treatment condition.

© Substitute Teacher
Figure 11. Mean percent of intervals of Appropriate behavior for S12 and S13 in Ms. A's class under each treatment condition.

© Substitute Teacher
Figure 12. Mean percent of intervals of Appropriate behavior for S14 and S15 in Ms. A's class under each treatment condition.

© Substitute Teacher
behind. Consequently, he was required to wait for long periods of time as the teacher instructed the other students; these periods without a work assignment may have accounted for the greater variability in his appropriate behavior.

Changes in individual categories of inappropriate behavior for Ms. A's class are shown in Table 3. Talking, as well as Time off Task behavior, each initially occurring 9.5 percent of the time, accounted for much of the maladaptive classroom behavior observed during the Baseline period. These responses were reduced during the Token+ condition to averages of 3.3 and 1.5 percent respectively. Slight decreases were also observed for Out of Seat, Orienting, and Noise responses with the token economy. As the previous data showed, however, there were no reliable differences between the Token+ and the Token− conditions. Also, as in both the Period 1 and study hall settings, Talking responses were most affected by the removal of the program.

In summary, in both EMR classes and for twelve of the fifteen Ss observed, the introduction of the token reinforcement program increased the occurrence of appropriate classroom behavior over the averages observed during the Baseline and Withdrawal periods. No differences, however, were observed as a function of the two Token conditions.
Table 3
Average Percentages for Specific Behavior Classes
For Each Experimental Phase in Ms. A's Class

<table>
<thead>
<tr>
<th>Behavior Classes</th>
<th>Baseline</th>
<th>Token+₁</th>
<th>Withdrawal</th>
<th>Token+₂</th>
<th>Token+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate</td>
<td>75.4</td>
<td>92.3</td>
<td>88.0</td>
<td>94.8</td>
<td>89.8</td>
</tr>
<tr>
<td>Time off Task</td>
<td>9.5</td>
<td>3.3</td>
<td>4.0</td>
<td>2.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Out of Seat</td>
<td>1.9</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>1.1</td>
</tr>
<tr>
<td>Orienting</td>
<td>4.0</td>
<td>1.8</td>
<td>1.8</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Noise</td>
<td>2.5</td>
<td>0.4</td>
<td>1.0</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Disturbing Others</td>
<td>0.5</td>
<td>0.4</td>
<td>1.0</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Talking</td>
<td>9.5</td>
<td>1.5</td>
<td>6.0</td>
<td>1.3</td>
<td>1.8</td>
</tr>
</tbody>
</table>
Comparison of Adaptive Behavior Levels in Normal and EMR Classrooms

Figure 13 shows the mean occurrence Appropriate behavior for each of the EMR classes during the first eight days of the Baseline condition and for the group of students in the normal class during a similar "baseline" period. There was some overlap between on-task behavior in Ms. A's class and that in the normal classroom, although the mean for the special education class was higher ($\overline{X} = 72.1$ for the EMRs versus $\overline{X} = 60.4$ for the normals). It was apparent for Ms. B's class, however, that even during the Baseline period, EMR students were more frequently engaged in task-relevant behavior (EMR $\overline{X} = 81.7$, and normal $\overline{X} = 60.4$).

An examination of the daily behavior of individuals in the normal class showed that one of the Ss, chosen from the group that the teacher had named as the most disruptive, failed to exhibit any appropriate behavior on any of the days on which he was observed. He either roamed about the room talking to other students or slept at his desk. When his data are eliminated, the means more accurately reflect the behavior of the other students in the normal sample (see Figure 14 for the adjusted curve in comparison to the EMR classes). With this adjustment the normal and the EMR students exhibited more similar levels of Appropriate behavior ($\overline{X} = 68.1$ percent for the normals; $\overline{X} = 72.1$ percent for Ms. A's class; $\overline{X} = 81.7$ percent for Ms. B's class).
Figure 13. Mean percent of intervals of Appropriate behavior in the EMR and normal classrooms during the Baseline condition.
Figure 14. Adjusted mean percent of intervals of appropriate behavior in the EMR and normal classrooms during the Baseline condition.
As shown in Table 4, Time off Task and Talking were the most frequently occurring inappropriate behaviors in both the EMR and normal classrooms. However, the means of each response were higher for the normal students. Time-off-Task behavior occurred an average of 18.9 percent of the time in the regular class, whereas EMR students exhibited non-task behavior in 12.2 percent of the intervals in Ms. A's class and in 10.8 percent in Ms. B's class during the Baseline period. Normal students were observed to talk without permission in 16.9 percent of the intervals; in the special education classes the students spent 9.7 percent of the observed time talking in Ms. A's room and 2.3 percent in Ms. B's room. Comparisons of the other inappropriate behavior measures yielded only small differences.

The addition of the token reinforcement program to the special education classes brought the mean percentage of Appropriate behavior well above that observed in the normal classroom. Figure 15 shows the average percent of intervals in which Appropriate behavior occurred in each of the EMR classes during the Token+ condition plotted against the average observed in the normal class during the Baseline condition. On-task behavior in both EMR classes (̅X = 94.8 in Ms. A's class and ̅X = 93.3 percent in Ms. B's room) was more frequent than that in the regular class for either all eight Ss in the normal sample (̅X = 60.4 percent) or for the
Table 4
Average Percentages for Specific Behavior Classes
For Normal and EMR classes in the Baseline Condition

<table>
<thead>
<tr>
<th>Behavior Class</th>
<th>EMR A</th>
<th>EMR B</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate</td>
<td>72.1</td>
<td>81.7</td>
<td>60.4</td>
</tr>
<tr>
<td>Time off Task</td>
<td>12.2</td>
<td>10.8</td>
<td>18.9</td>
</tr>
<tr>
<td>Out of Seat</td>
<td>1.4</td>
<td>0.1</td>
<td>5.2</td>
</tr>
<tr>
<td>Orienting</td>
<td>4.2</td>
<td>1.6</td>
<td>4.9</td>
</tr>
<tr>
<td>Noise</td>
<td>1.8</td>
<td>2.2</td>
<td>1.0</td>
</tr>
<tr>
<td>Disturbing Others</td>
<td>0.6</td>
<td>0.7</td>
<td>3.3</td>
</tr>
<tr>
<td>Talking</td>
<td>9.7</td>
<td>2.3</td>
<td>16.9</td>
</tr>
</tbody>
</table>
Figure 15. Mean percent of intervals of Appropriate behavior in EMR classrooms with token reinforcement and a normal class in the Baseline condition.
adjusted group (X̄ = 68.1 percent). It is also interesting to note that the EMR students whose behavior was least affected by the token program still displayed more task-relevant responses than did the most disruptive student in the normal class.

In summary, it appears that a token reinforcement program in the special education classes in this junior high school increased the adaptive behavior of the EMR students in terms of appropriate classroom behavior, as compared to the behavior of children in normal classrooms. And perhaps of even more interest, it should be noted that this increase in the percentage of time spent in appropriate behaviors came on top of an already higher level of on-task behavior in the EMR classes than the normal classroom when they were compared during the initial baseline periods.
Chapter IV
DISCUSSION

It is apparent that the implementation of a token reinforcement program in a special education classroom is effective in raising and maintaining very high levels of appropriate social behavior among EMR students, even when their initial base rates of appropriate behavior are quite high, in this particular school setting. However, manipulation of the token-behavior arrangement in terms of imposing positive reinforcement or positive reinforcement plus response costs has little differential effect on the frequency of occurrence of adequate adaptive behavior in the classroom.

Although there was no empirical evidence upon which to base such an assumption, it had been expected that with EMR students in classroom settings, positive reinforcement plus response cost procedures would be more effective than positive reinforcement procedures alone because response costs might serve to emphasize the existence of behavior-reinforcement relationships, one function of token economies suggested by Vernon (1972), by additional concrete cueing of negatively sanctioned behaviors as well as acceptable behaviors.

Several factors may have accounted for this non-differential effect of the procedures. First, it may
have been that the aversive consequences, i.e., the withdrawal of tokens, did not occur at the most effective time in the behavior sequence. Laboratory studies would suggest that response inhibition is greatest when the punishing stimulus occurs immediately after the negatively sanctioned behavior. In the present investigation, this was not the case in most instances. During the first weeks of the study, the teachers gave a warning and then a quiet announcement that the individual breaking the rules would lose points. The actual removal, however, occurred at the end of class, which in some cases was as much as thirty or forty minutes after the negatively sanctioned behavior occurred. A more immediate consequence, such as points being removed from the offender's card on the spot or the delivery of tickets of some sort payable at the end of the class hour might have been more effective in reducing an undesirable response. In the classes under investigation, there was a great deal of sniggering and interruption in on-task behavior when a student received a reprimand. Because of this observation and reports from other investigators that obvious reprimands can actually increase disruptive behavior (e.g., Thomas, et al., 1968), an immediate, highly visible means of removing tokens was not used.

Second, although the experimenter made every effort to stress the importance of adhering to the guidelines of the conditions, and specifically that points were not to be
removed from previously earned totals during the positive only phase, Ms. B continued to levy occasional fines regardless of condition. Thus, the subjects from her class may not have been able to discriminate between the two token arrangements at all times. Since the students changed from room to room throughout the day, it was not possible to resolve this problem by examining data from Ms. A's classes only.

The third, and as important as any, reason for no apparent differences between the two token procedures was that there may well have been a ceiling effect on the occurrence of the desired behaviors. In conversations with the teachers prior to actual observation, both reported frequent and serious disruptive acts on the part of several pupils. Observation, however, failed to confirm their reports. Even before token conditions were imposed, the students in both classes exhibited higher frequencies of appropriate behavior than has been reported for normal classes in this same school and that reported by other investigators who have developed classroom token programs. In fact, the Baseline means in the EMR classes in this study correspond to averages obtained during token conditions in a number of other studies (e.g., Broden, et al., 1970; O'Leary, et al., 1969). Therefore, this restricted range could have contributed to the lack of differentiation in the assessed effectiveness of the two token systems.
There were, however, some interesting differences in the students' affective reactions to the program. They were not polled systematically, but comments that the observers heard in class and reports from the teachers indicated that the Token+ condition was better liked. When the teachers crossed out previously earned points in the response cost condition, there were frequent exclamations of dismay. On at least two occasions in Ms. B's class, students tore up their cards following the removal of points. It is unclear whether this action was the result of the response cost procedure itself or that Ms. B had unfairly removed the points. A third student, cited before the study began as the most disruptive in the class but not observed because he was isolated by partitions, became so upset over perceived inequitable point losses that he was sent to the office and ultimately suspended for five days for the disturbance he caused. O'Leary and Drabman (1971) warned that a response cost contingency might increase the punitiveness of the reinforcing agent and thus undermine the effectiveness of the token economy. It seems likely, however, that such an effect depends upon the individual teacher's baseline approach to discipline, since no incidents of this type were reported by Ms. A who was less punitive initially.

While one or the other variant of the token procedure did indeed increase task-relevant behavior and decrease
time-off-task and talking (the traditional measures of undesirable classroom behavior), there were still occurrences of maladaptive behavior deemed serious enough by the teachers that they insisted on the removal of two target subjects from the program—and thus the token economy could not be considered a complete success. S2, furthermore, was suspended three and one-half weeks before the end of the term after he "talked back" to Ms. B and earned enough days of suspension to qualify for exclusion for the remainder of the year. Another student, though not a target subject, whose behavior was much improved under the Token+ condition refused to cooperate when the token economy was reinstated following the Withdrawal period.

Some adjustments in the token system might have salvaged these students. Well into the Token+ condition, the experimenter learned from a casual comment of one of the students that one teacher was allowing her students to buy only one item from the store each day on which points were exchanged. Consequently, several of the students were accumulating many more points than they needed to buy any of the available items. There was no need therefore to meet the requirements to earn more points. Some arrangement which requires the student to relinquish all or most of his points at each redemption period is no doubt a more effective procedure.
Behavior modifiers who have employed token reinforcement systems have been criticized for their apparently greater interest in classroom control than in improving academic performance. Originally, this study was to have examined the effects of token reinforcement on academic performance in an on-going classroom. Most of the studies focusing on achievement have involved curricula of programmed instruction in settings such as university laboratory schools where many of the variables could be controlled. Unfortunately, in the present classroom settings which were to have constituted a departure, there were no quantifiable academic behaviors adaptable to token contingencies because individual academic performance was obviated by the teachers' preemptive procedures. For example, in Ms. B's class, the lessons, either from social studies workbooks or the weekly newspaper, consisted of her putting the correct answers on the board and the students copying them down. In Ms. A's reading class, she worked with a small group while the rest of the pupils were given desk work (e.g., crossword puzzles) which was gone over in class and corrected as a group before it was handed in. In the math class the students worked through a series of multiplication drills. Most had done the same series the previous year, and the baseline rates of accuracy were so high that no improvement could have been effected by the imposition of token contingencies. Thus, what was to have been an investigation focusing on both
academic behavior in an essentially non-programmed classroom setting and on adaptive behavior in the same setting was restricted to the latter only. However, since there is much current interest in the notion of impairment in adaptive behavior and since it would appear to be the more malleable variable (relative to intellectual functioning) in AAMD's definition of mental retardation, the social focus is not without justification and, hopefully, not without reinforcement value in and of itself.

Perhaps the most significant finding in the present investigation, given the current social and legal milieu in the state of California (i.e., that impairment in adaptive behavior in addition to subaverage intelligence must be demonstrated for inclusion in EMR classes), is that the comparison of appropriate behavior in normal and special education classes in a school system in California has revealed that even prior to the implementation of the token program, EMR students exhibited a higher level of acceptable behavior than their normal counterparts. And it is worth re-emphasizing that after the introduction of the token systems, appropriate behavior for the majority of the EMR students increased and remained at exceedingly high levels for the duration of the study.

Even if it were to be granted that these classes of EMRs and normals and their teachers might be atypical, the fact remains that most of the students in these EMR classes
did not show an impairment in adaptive behavior when they were compared with their normal peers. Thus, it would not seem appropriate for them to have been labeled mentally retarded as based upon the adaptive behavior portion of the AAMD standard.

Unfortunately, at this point in their academic careers in public school, it would appear impossible to ascertain whether these so-called EMR children actually exhibited maladaptive social behaviors which led to their being assigned the label or whether they were, in fact, separated solely on the basis of their intelligence test scores (rather than a documented combination of both criteria) very early-on. As is self-evident, by the time they have reached junior high school, and given the time wasted and the curriculum restrictions endured in special classes, reassignment to regular classes has become virtually impossible—regardless of the effectiveness and efficiency of token reinforcement procedures in shaping up their socially adaptive classroom behaviors.

However, on the optimistic side, if the kinds of procedures which have been effected in the present study with junior high school aged students can be shown to lead to high levels of appropriate social behaviors, it would seem reasonable to suggest that these same kinds of procedures might be much more effective and probably more easily implemented with younger children who might be
expected to have less variegated behavioral repertoires—before they too have been victimized by teacher expectancies and similar labeling processes.

Of course, the ultimate test of the usefulness of the token economy as a remedial tool in any situation is whether behavior which was shaped using token reinforcement can be maintained when the token support is withdrawn. Unfortunately in the present study, the semester ended before there was an opportunity to attempt to fade out the program. Presumably, reinforcers occurring naturally in the school environment, such as cooking privileges, access to an art corner, or free time could have been substituted gradually for the material reinforcement. At the same time the periods between opportunities to exchange points for backup reinforcement could have been lengthened along with raising the standards of behavior necessary to earn points. Moreover, withdrawal of the token economy without concomitant decreases in appropriate behavior may be more successful if the students themselves are first given the responsibility of keeping daily behavior records and actually awarding points to themselves. Self-reward has not been systematically investigated by researchers interested in contingency management, and it may prove to be an important aspect of the internalization of behavioral standards.
The present study sought to examine the relative effectiveness of two token-behavior arrangements in a token economy—positive reinforcement only and positive reinforcement plus response costs—in modifying maladaptive classroom behavior in two special education classes in a junior high school. It was expected that with EMR students in classroom settings, positive reinforcement plus response cost procedures would be more effective than positive reinforcement procedures alone because response costs might serve to emphasize the existence of behavior-reinforcement relationships by additional cueing of negatively sanctioned behaviors as well as acceptable responses.

Contrary to expectation, while the introduction of the token reinforcement program increased the occurrence of appropriate classroom behavior over the averages observed during the Baseline period, no differences were observed as a function of the two Token conditions. Possible reasons for the non-differential effects were discussed, especially with regard to the general problem of timing of response cost delivery and to several problems particular to this study.
A second aim was to assess the relative effectiveness of both of these token schemes in building up and maintaining a level of appropriate adaptive behavior in EMR classrooms comparable to that expected from students in a "normal" classroom in the same school. It was found that the introduction of a token reinforcement system in the special education classes in this school increased appropriate behavior among the EMR students well above that tolerated in the regular class, and in fact the increase came on top of an already higher level of on-task behavior in the special education classes when compared with the normal classroom during the baseline periods.

The implications of these findings for the usefulness of token economies as a remedial tool in EMR students' retaining or regaining regular class placement were noted, as well as suggestions for fading out the token program.
<table>
<thead>
<tr>
<th>Points</th>
<th>Rule Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+2</td>
<td>Stay in your seat.</td>
</tr>
<tr>
<td>+2</td>
<td>Keep quiet during class. No talking or noise.</td>
</tr>
<tr>
<td>+2</td>
<td>Raise your hand before speaking.</td>
</tr>
<tr>
<td>+2</td>
<td>Leave other people and their things alone.</td>
</tr>
<tr>
<td>+2</td>
<td>Be courteous in the halls between classes.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Points</th>
<th>Rule Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>Being out of your seat when not allowed.</td>
</tr>
<tr>
<td>-2</td>
<td>Talking and noise during class.</td>
</tr>
<tr>
<td>-2</td>
<td>Speaking out without raising your hand.</td>
</tr>
<tr>
<td>-2</td>
<td>Bothering other people and their things.</td>
</tr>
<tr>
<td>-2</td>
<td>Poor hall behavior.</td>
</tr>
<tr>
<td>-2</td>
<td>Talking back to the teacher.</td>
</tr>
</tbody>
</table>
APPENDIX B

Revised Class Rules

Each period you may earn points for following these class rules:

Points

+2 Stay in your seat.
+2 Keep quiet during class.
+2 Leave other people and their things alone.
+2 Have all your supplies for class.
+2 Work the whole period.

If you earn all 10 points for the period, you will receive a bonus of 5 points.
REFERENCES


Clark, M., Lachowicz, J., and Wolf, M. A pilot basic education program for school dropouts incorporating a token reinforcement system. Behaviour Research and Therapy, 1968, 6, 183-188.


----- Behavior Modification in Mental Retardation. Chicago: Aldine-Atherton, 1971


Hall, R., Axelrod, S., Tyler, L., Grief, E., Jones, F., and Robertson, R. Modification of behavior problems in the home with a parent as observer and experimenter. Journal of Applied Behavior Analysis, 1972, 4, 53-64.


