INFORMATION TO USERS

This material 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.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeib Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR
LOZANO, Servando Ruben, 1940-
THE EFFECTS OF VISUAL CUEING AND AUDITORY
CUEING ON SELF-CONTROL IN CHILDREN.

The Ohio State University, Ph.D., 1976
Education, special

Xerox University Microfilms, Ann Arbor, Michigan 48106

© 1977

SERVANDO RUBEN LOZANO

ALL RIGHTS RESERVED
THE EFFECTS OF VISUAL CUEING AND AUDITORY CUEING
ON SELF-CONTROL IN CHILDREN

DISSERTATION

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

By
Servando Ruben Lozano, B.A., M. Ed.

* * * * *

The Ohio State University
1976

Reading Committee:
Dr. John Cooper
Dr. Donald Tosi
Dr. Thomas Stephens

Approved By
John E. Cooper
Adviser
Faculty for Exceptional Children
ACKNOWLEDGMENTS

Tenacity may be the single most attribute necessary to complete a dissertation. Dr. Charles B. Huelsman, Jr. encouraged and stimulated my waning efforts. I deeply regret that Dr. Huelsman's sudden death ended my association with a good man and an excellent teacher.

The direction and timely assistance by Dr. John Cooper, my adviser, was of crucial importance. His perceptive analysis and ready availability made this dissertation possible.

The urging, support and timely advice provided by my wife, Martha, gave the efforts toward completion of the dissertation meaning. The satisfaction of finishing is sweet when I can share it with her.

There is no way to adequately express my indebtedness to my family in general and my mother in particular. The justification of the effort and the confidence in the outcome are due in large part to their consistent reassurance.

The vague threats by my mother-in-law helped too.

Eulah Grimm deserves a special thanks for her cooperation and assistance. A special thanks to Gina Shar and Steve King.
VITA

December 31, 1940 .......... Born, San Marcos, Texas

1967 ...................... B.A., Bowling Green State University, Bowling Green, Ohio

1968-1969 ................. Intern-School Psychologist, Findlay City Schools, Findlay, Ohio

1969 ...................... M. Ed., Bowling Green State University, Bowling Green, Ohio

1969-1971 ................. School Psychologist, Southwestern City Schools, Grove City, Ohio

1971-1973 ................. Behavioral Consultant, Research and Demonstration Unit, Southwestern City Schools, Grove City, Ohio

1973-1974 ................. Fellowship, Faculty for Exceptional Children, The Ohio State University, Columbus, Ohio

1974-1975 ................. Graduate Research Associate, Faculty for Exceptional Children, The Ohio State University, Columbus, Ohio


FIELD OF STUDY

Major Field: School Psychology
# 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>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
</tbody>
</table>

## Chapter

### I. INTRODUCTION .................................... 1
- Background ........................................ 1
- Purpose .......................................... 7
- Scope of the Study ................................ 7
- Limitations of the Study ......................... 9
- Questions of the Study ........................... 10
- Definitions of Terms ............................. 11

### II. REVIEW OF THE LITERATURE ....................... 13
- Self-Recording .................................... 14
- Cognitive Mediation and Self-Control .......... 19
- Laboratory Studies in Self-Control .......... 41
- Self-Control Studies in Applied Settings .... 81
- Summary ......................................... 114

### III. METHODS AND PROCEDURES ........................ 117
- Subjects ........................................ 117
- Setting ......................................... 118
- Materials ...................................... 119
- Behavior Definition and Recording .......... 120
- Design of the Study ............................ 123
- Design Conditions ................................ 125
- Data Collection and Analysis .................. 131
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV. PRESENTATION AND ANALYSIS OF DATA</td>
<td>134</td>
</tr>
<tr>
<td>Inter-Observer Agreement</td>
<td>134</td>
</tr>
<tr>
<td>Presentation and Analysis of Data</td>
<td>137</td>
</tr>
<tr>
<td>Group Results</td>
<td>151</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>156</td>
</tr>
<tr>
<td>Discussion of Data</td>
<td>156</td>
</tr>
<tr>
<td>V. SUMMARY, DISCUSSION AND RECOMMENDATIONS</td>
<td>162</td>
</tr>
<tr>
<td>Summary</td>
<td>162</td>
</tr>
<tr>
<td>Discussion</td>
<td>164</td>
</tr>
<tr>
<td>Recommendations</td>
<td>174</td>
</tr>
<tr>
<td>VI. REFERENCES</td>
<td>177</td>
</tr>
<tr>
<td>Appendix A</td>
<td>177</td>
</tr>
<tr>
<td>Appendix B</td>
<td>179</td>
</tr>
<tr>
<td>Appendix C</td>
<td>181</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>183</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Scored and Unscored Interval Percentages of Agreement, Raw Data,</td>
<td>135</td>
</tr>
<tr>
<td>and Total Percentages of Agreement, by Condition, for On-Task Behavior</td>
<td></td>
</tr>
<tr>
<td>2. Scored and Unscored Interval Percentages of Agreement, Number of</td>
<td>135</td>
</tr>
<tr>
<td>Measures, and Total Percentages of Agreement, by Condition, for Teacher Praise</td>
<td></td>
</tr>
<tr>
<td>3. Scored and Unscored Interval Percentages of Agreement, Number of</td>
<td>136</td>
</tr>
<tr>
<td>Measures, and Total Percentages of Agreement, by Condition, for Accurate Self-Reinforcement</td>
<td></td>
</tr>
<tr>
<td>4. Scored and Unscored Interval Percentages of Agreement, Number of</td>
<td>136</td>
</tr>
<tr>
<td>Measures, and Total Percentages of Agreement, by Condition, for Accurate Self-Punishment</td>
<td></td>
</tr>
<tr>
<td>5. Raw Data and Agreement Percentages Between Individuals and the</td>
<td>152</td>
</tr>
<tr>
<td>Observer for Self-Reward and Self-Punishments, by Condition</td>
<td></td>
</tr>
<tr>
<td>6. Values for l-g Obtained in the Distribution-Free Multiple</td>
<td>157</td>
</tr>
<tr>
<td>Comparisons Based on Friedman Rank Sums</td>
<td></td>
</tr>
<tr>
<td>7. The Percentage of Intervals On-Task, Total Accuracy Percentage</td>
<td>171</td>
</tr>
<tr>
<td>of Self-Reward and Percentage of Reinforcement Self-Awarded During Self-Reinforcement, per Individual</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Representation of the Reversal Design Employed in the Study</td>
<td>126</td>
</tr>
<tr>
<td>Plate</td>
<td></td>
</tr>
<tr>
<td>1. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject One</td>
<td>139</td>
</tr>
<tr>
<td>2. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject Two</td>
<td>141</td>
</tr>
<tr>
<td>3. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject Three</td>
<td>143</td>
</tr>
<tr>
<td>4. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject Four</td>
<td>145</td>
</tr>
<tr>
<td>5. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject Five</td>
<td>147</td>
</tr>
<tr>
<td>6. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject Six</td>
<td>149</td>
</tr>
<tr>
<td>7. The Daily and Total Percentages of Intervals, On-Task, Per Condition for Subject Seven</td>
<td>150</td>
</tr>
<tr>
<td>Figure</td>
<td></td>
</tr>
<tr>
<td>2. The Daily Frequency of Teacher Praise Per Condition and The Daily Percentages of Intervals On-Task for The Group by Condition</td>
<td>154</td>
</tr>
<tr>
<td>3. The Daily and Total Percentages of Intervals On-Task for All Seven Children Per Condition</td>
<td>155</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

The main thrust of behavioral research has been the investigation of functional relationships between behavior and environmental variables: external control. A great deal of research has resulted from this emphasis on external control, especially investigations with infrahuman subjects.

Theoretically, much of this research has been influenced by Watsonian behaviorism (1919) and the emphasis by Skinner (1953) on environmental determinants of behavior. Little attention to self-control has been a result. Perhaps another reason self-directed behavior change has received little attention was the unscientific status of self-observation of cognitive events. The importance of self-control has not been devalued by all researchers. Skinner (1953) asserted that behavior can be maintained in the absence of external reinforcement. Also, Bandura (1969) stated:

There exists ample evidence that one cannot account satisfactorily for human behavior while remaining entirely outside the organism, because overt behavior is often governed by self-generated stimulation that is relatively independent of environmental stimulus events (p. 39).

Since the early 1960's, an increasing amount of research has been published regarding self-control. Pioneering work by Kanfer and Marston investigated the degree of learning and accurate self-reinforcement...
(Kanfer, Bradley and Marston, 1962), the degree of learning and rate of self-reinforcement (Kanfer and Marston, 1963a) and the rate of self-reinforcement and social reinforcement (Kanfer and Marston, 1963b). Adults were the focus of the early research, but investigators soon turned to children as the acquisition of self-control was analyzed.

The imitation of self-reward standards by children were soon examined by Bandura and his associates (Bandura and Kupers, 1964; Bandura and Whalen, 1966). A series of experiments by Masters investigated the influence of social comparison processes on children's self-rewarding behavior (Masters, 1968, 1969, 1971 and 1973). Kanfer and associates also studied self-rewarding by children. The focus of much of Kanfer's research was the influence of subject variables such as age, class standing, and level of incentive (Kanfer, 1966). A great deal of research in self-control has followed with the bulk of the studies implemented in the laboratory.

Recently, a number of investigations in self-control have been implemented in classrooms or other applied settings. Some of these studies have demonstrated surprising effectiveness after short treatment periods. For example, after one hour of instruction two mothers self-recorded attention to appropriate behavior. The mothers' self-recording attention to appropriate behavior dramatically increased appropriate behavior (Herbert and Baer, 1972). Also, swimming coaches very quickly eliminated most or all of several behaviors incompatible with effective swimming performance by having swimmers publicly self-record the frequency of the undesirable behaviors (McKenzie and Rushall,
1974). A different approach was found effective in reducing rule infractions in school. Students who repeatedly broke specific rules were exposed to information designed to supplant cognitive awareness of the task and reduce a cognitive deficit regarding solutions. This cognitive mediational approach was effective with chronic rule breakers (Blackwood, 1970) and children causing lunchroom disturbances (MacPherson, Benjamin and Hohman, 1974).

Several of the applied studies have compared the effectiveness of self-reward to external reward with different student populations. Self-reinforcement was found to be at least as effective as external-reinforcement with regular class children (Felixbrod, Jeffrey and O'Leary, 1973; Glynn, 1970), with disruptive regular class children (Bolstad and Johnson, 1972), and with learning disabled children (Lovitt and Curtiss, 1969). Self-reward was effective in maintaining pro-social behaviors with adolescent boys residing in a psychiatric hospital (Kaufman and O'Leary, 1972).

Classroom studies have demonstrated increased on-task behavior (Glynn and Thomas, 1974; Glynn, Thomas and Shee, 1973), decreased disruptive behavior (Bolstad and Johnson, 1972; Drabman, Spitalnik and O'Leary, 1972) and improved academic performance (Felixbrod and O'Leary, 1973; Glynn, 1970; Lovitt, 1970; Lovitt and Curtiss, 1973).

The acquisition of self-reward behavior has been of major interest since the first study of self-control in children (Bandura and Kupers, 1964). The applied studies which incorporated self-reward training began by externally-rewarding the children, then training the children to self-reward and lastly allowing the children to self-reward without
external intervention. Relatively short periods of external-reward and training have been effective with regular class children. Disruptive first and second graders were trained to self-reward after thirteen days of experimenter reinforcement and training (Bolstad and Johnson, 1972). The training used was reinforcement contingent upon accurate self-evaluation and self-recording of performance. The same training procedure was used to teach self-rewarding to a public school class for children with behavioral difficulties (Drabman, Spitalnik and O'Leary, 1973). The procedure proved effective, but the training period was thirty days and was preceded by five days of external-reinforcement. Kaufman and O'Leary (1972) were successful in training adolescent psychiatric residents to self-reward after twenty-five days of external-reward. When a similar transition was tried with the same population after only nine days of external-reinforcement, the self-reward procedure was ineffective (Santogrossi, O'Leary, Romanczyk and Kaufnan, 1973). To the extent that self-reward, an effective self-control procedure, must be preceded by lengthy training, the practicality and cost/benefit of implementation are seriously reduced. A procedure to implement self-reinforcement without any training would be much more practical and relatively inexpensive.

Such a procedure was developed by Glynn, Thomas and Shee (1973). The refined procedure included auditory cues, at random time intervals, which signaled the opportunity to self-evaluate performance and then self-reward. The children marked a card on their desk when cued, if assigned work was being performed. When this procedure was implemented
with third graders in regular class, the techniques did not work well (Glynn and Thomas, 1974). The children seemed unsure what task was assigned. For example, the teacher sometimes asked for attention after initial instructions to give additional directions. Those children who ignored the teacher and continued to work at their desks were actually off-task, but they considered themselves on-task.

Another factor seemed to be the frequency of the auditory cues. The auditory cues occurred after random intervals of one, two, three, four or five minutes. Observers noted that the four and five minute intervals seemed too long to maintain the on-task behavior. As a result of the authors' analysis, the procedure was modified in two ways:

1. A visual cue (a sign) was placed at the front of the room to indicate what behavior was appropriate, attending to the teacher or working at the desk;

2. The longer intervals were removed leaving only one, two and three minute intervals between auditory cues.

Since a reversal design was in use, the modified procedure was implemented after a second baseline and proved to be highly effective. The authors had reached their objective of increasing the effectiveness of the cueing procedure. Also, for the first time a published study established effective self-reward with no training. But the study left several unresolved issues. The methodological faults and confounding of variables seriously limited the number of tenable inferences.
The most serious methodological fault with the study was in the incompleteness. A reversal design was used. With this design, causal inferences can only follow when the experimental procedure has been introduced twice, i.e., the experimental procedure must be introduced subsequent to a baseline and again after a second baseline (Cooper, 1974). The sequence of procedures would be as follows:

1. Baseline;
2. Experimental procedure;
3. Baseline again;
4. Experimental procedure, again.

When the undesirable behavior increases during the second baseline and then decreases a second time when the same experimental procedure is re-introduced, a causal or functional relationship between the experimental procedure and the decrease in undesirable behavior has been cogently demonstrated. Adding a visual cue and altering the length of the intervals between auditory cues resulted in a new procedure. This new procedure was not removed and then re-introduced. Also, by eliminating the four and five minute intervals, the opportunity to self-reward was increased by 50%. This apparent increased self-reinforcement may have resulted in the new procedure being more powerful. The effect of the visual cue is uninterpretable due to the confounding with the altered time intervals. The resultant confounding of variables and the incompleteness of the reversal design make it impossible to infer functional relationships between the modified experimental procedure and the obtained behavior change.
Purpose

The purpose of this study was to evaluate the effectiveness of the auditory and visual cueing procedures utilized by Glynn and Thomas (1974) to establish self-control. The study maintained equal opportunity for reinforcement during experimental treatments and the effects of visual and auditory cueing were evaluated independently and in combination. The most effective of the three treatments was implemented again in congruence with a reversal design to more adequately investigate the effects. Glynn and Thomas (1974) investigated the effects of the self-control procedure with regular third graders. Learning disabled children of first through second grade age were utilized in this study to evaluate the applicability of the procedure with this population.

Scope of the Study

The study evaluated the effect of visual cueing, auditory cueing with self-reward, and visual and auditory cueing with self-reward upon the on-task behavior of learning disabled students. Also evaluated was the accuracy of the student's self-reinforcement and self-punishment. An accuracy score was obtained by computing the percentage of matches between the student's self-evaluation of being on-task and an observer's evaluation of the student's on-task behavior. The frequency of teacher praise was also measured to assess whether the visual or auditory cueing elicited different amounts of teacher praise.
Six different design phases were utilized in the study. Following baseline, during which on-task and teacher praise were measured, visual cueing (Intervention 1) was introduced. Auditory cueing and self-reward (Intervention 2) was introduced next. During the interventions involving self-reward, the accuracy of the subjects' self-evaluations was measured in addition to on-task behavior and teacher praise. The next phase was visual cueing combined with auditory cueing and self-reward (Intervention 3). Baseline 2 was introduced next preceding the final phase, the implementation of the most effective of the three interventions implemented. Intervention 3 was the most effective and was re-introduced.

Seven of ten students in a learning disabilities/behavior disorders (LD/BD) class were selected as the experimental population. On-task behavior occurred during 75% or less of the experimental period for the primary level students selected.

The experimental period was the same daily forty-five minute period of reading instruction. During the interventions involving self-reward, the self-awarded x's were exchanged for rewarding activities or objects during the last thirty minutes of the school day.

An observer recorded on-task behavior and, when appropriate, self-rewarding behaviors, using a sequential time-sampling procedure. Teacher praise was continuously recorded. A reversal design was utilized to analyze the interventions. During each experimental phase, the performance of each subject was measured. Each experimental phase was maintained until stabilization occurred. In addition to the
reversal design analysis, the on-task data were subjected to a non-parametric statistical analysis for the effects of the experimental analysis.

**Limitations of the Study**

The study was designed to assess the effect of visual cueing, auditory cueing with self-reward, and visual and auditory cueing with self-reward upon the on-task behavior of seven learning disabled, primary age students.

Due to the limited number of children employed in the study and due to the method of selecting the children no generalizations can be made to children other than those in this study.

This study investigated the relationship between on-task behavior and the auditory and visual cueing of self-reinforcement behaviors. The effects of visual cueing and auditory cueing of the self-control of on-task behavior were evaluated independently. The visual cue was implemented alone, the auditory cue was used alone, then both cues were used together to evaluate the separate effects. The opportunity to self-reinforce was maintained at equal levels among the phases utilizing self-rewards, to eliminate a confounding effect. The most effective of the three procedures, visual cueing, auditory cueing with self-reward, or visual and auditory cueing with self-reward, was reintroduced within a reversal design to allow inferences regarding causality. Also assessed was the accuracy of the students' self-evaluations. Teacher praise was measured during the experimental phases.
to insure that the cueing was not influencing the frequency of teacher praise.

Questions of the Study

A learning disabilities/behavior disorders class was utilized in this study, therefore the objectives pertain only to the selected class. The study gathered evidence to answer the following questions:

1) Is the combined visual cueing and auditory cueing with self-reinforcement, effective in increasing the on-task behavior of students in a learning disabilities class?

2) Is the effect of visual cueing alone, different from the effect of auditory cueing with self-reinforcement?

3) Are the effects of auditory cueing with self-reward and auditory cueing with self-reinforcement different?

4) Is the students' self-reward for on-task behavior and self-punishment for off-task behavior accurate, i.e., does the observer's evaluation match the child's self-recording?

5) Is the teacher's praise unaffected by the cueing procedures?

The study investigated the following null hypothesis:

1) There will be no statistically significant differences among the experimental conditions involving visual cueing, auditory cueing with self-reward, visual and auditory cueing with self-reward, baseline and baseline 2.
Definitions of Terms

1. **Self-control**: In the relative absence of immediate external constraints, a person's engaging in behavior of lower probability than that of alternatively available behaviors (Thoresen and Mahoney, 1974).

2. **Self-recording**: A child's recording the occurrence of a specific behavior of his own.

3. **Self-reinforcement or self-reward**: The self-evaluation that a specific behavior has occurred, and the self-awarding of points exchangeable for a reinforcer.

4. **Back-up**: An object or activity which, when exchanged for points that are contingent upon a specific behavior, acts as a reinforcer for that particular behavior.

5. **Visual cue**: A discriminative stimulus which indicates which class of behaviors are appropriate for self-reinforcement.

6. **Auditory cue**: A sound which signals when self-evaluation should occur to judge whether self-reinforcement is appropriate.

7. **Learning disabilities/behavior disorders (LD/BD) class**: A special education class funded by the Division of Special Education in Ohio for children of elementary school age with I.Q.'s above 80, who have significant learning and/or behavior problems.

8. **Functional relationship**: The dependent variable and a given procedure are functionally related, if the behavior systematically varies as a function of the application of the procedure.

9. **Statistically significant difference**: Scores subjected to statistical analysis at the .05 level of confidence and found to be significant.
10. **Stabilization**: Data that are neither increasing or decreasing in frequency or data that are changing in frequency opposite to that obtained during the antecedent experimental phase.

11. **External-reinforcement**: Reinforcement or tokens contingent upon specific behavior and delivered at the sole discretion of somebody other than the receiver of the reinforcement.

12. **On-task behavior**: Performing assigned work or looking at the teacher when oral instructions are being given.

13. **Praise**: Oral statements encouraging or lauding behavior.

14. **Phase**: A period of time during which an experimental treatment is implemented.

15. **Sequential time-sampling**: Following a pre-selected sequence, each student is observed, consecutively for a specific time period. At the end of each time period, the occurrence or non-occurrence of pre-selected behaviors are recorded for that particular student.

16. **Baseline**: A series of consecutive experimental sessions during which dependent measures are obtained in the absence of an experimental treatment.

17. **Self-punishment**: The self-evaluation that a specific behavior has not occurred, and subsequently not self-awarding points exchangeable for a reinforcer.
CHAPTER II
REVIEW OF RELEVANT LITERATURE

Research in self-control has become increasingly prominent. Several authors have evaluated factors in self-control which in and of themselves comprise only a part of self-control, e.g., self-recording. The cognitive aspects of self-control have received much attention recently. Laboratory research has analyzed self-control from many viewpoints, but only recently have applied studies begun to assess the usefulness of the previously identified mechanisms of self-control. A review of relevant literature may clarify issues in the present study.

Self-control has meant many things to many people. The range of definitions was well documented by Kanfer (1970) and Mahoney and Thoresen (1974). Usually self-management, self-regulation, self-control, self-controlling behavior and self-controlled behaviors are not well delineated. For the purposes of this review, research was not selected on the basis of the definition of self-control utilized by the researchers, but by the inclusion of dependent or independent variables dealing with the following:

1) Self-determined consequences: self-selecting a schedule for consequences or a consequence, or both, for a behavior or class of behaviors;
2) Self-recording: recording the frequency of the performance of one's behaviors or class of behaviors;

3) Self-evaluation: judging whether one has performed a class of behaviors or a specific behavior;

4) Influencing or measuring cognitive variables to ascertain their influence on altering behavior probabilities.

The purpose of this review is to report the substantive findings of self-control research with public school age children. The scope of the review encompasses published research from 1965 to 1975, but it is not exhaustive. The majority of the published studies are included. When parametric or nonparametric statistical analyses were utilized, only statistically significant results of interest are reported. Relevant results are reported from those studies which utilized applied behavioral designs.

The research studies have been arbitrarily divided into four sections:

Self-recording;
Cognitive mediation and self-control;
Laboratory studies in self-control;
Self-control studies in applied settings.

Self-Recording

There are relatively few studies dealing only with the effects of self-recording, but those reported describe a range of different behaviors across different populations. In an often quoted study, Broden, Hall, and Mitts (1971) reported the work of a school counselor
who utilized self-recording in two situations: to increase the on-task behavior of an eighth grade girl, and to decrease inappropriate talk-outs by an eighth grade boy. By the use of applied behavior analysis, a reversal design and a multiple baseline design, the authors demonstrated self-recording to be effective, even though there was much variability in the correspondence between the observer's behavior recordings and the students' self-monitored behavior counts.

In a very similar study, Lovitt (1973) described using an applied behavior analysis design with two elementary pupils. Though different responses were problematic, hitting and talking out, the students quickly decreased the frequency of the undesirable behaviors.

The correspondence between self-reported behavior and actual performance was investigated by Risely and Hart (1968). The authors were able to significantly increase the accuracy of the self-reports of preschool children. Low income Negro children were praised and given snacks, if they accurately described their activities during a free play period. By calling attention to the behavior to be rewarded, e.g., "Did you paint today?" and then rewarding those who accurately reported painting, the teachers were able to increase painting behavior the following day. To gain the reward, the child had to remember to play with the appropriate materials the next day, and then self-report the activity. Those who inaccurately self-reported were told "No, you didn't paint today, did you." Those who accurately self-reported were reinforced. Play behavior was significantly altered. Though self-report does not meet the exact definition of
self-recording, it does involve many of the same activities, e.g., increasing the awareness of a behavior and recalling the occurrence of the behavior.

The problem of reliability of self-recording was investigated by Fixen, Phillips, and Wolf (1972). The authors evaluated the accuracy of self-recording by pre-delinquent boys, 12 to 15 years old, compared to peer-recordings and adult-recording. Though room cleanliness was carefully defined, the accuracy of self-recording was found to be unreliable. The youths overestimated the cleanliness observed. A reliability training procedure was evaluated and found to improve the correspondence between self- and peer-reporting, 76% agreement, but these reports were, as before, only 50% reliable compared to the adult observer. More stringent training regimes produced higher correspondences between adult-, peer-, and self-recordings: 73%, 79%, and 78%, respectively. Interestingly, self-recording had no significant effect on room cleanliness.

The accuracy or inaccuracy of self-recording had a quite different result in a study by Herbert and Baer (1972). After a one hour contact session, a short set of written instructions were given to the mothers of two difficult-to-manage children. Both mothers were instructed to self-record attention to appropriate behavior with wrist counters for two hours a day, five days a week. The accuracy of the correspondence between an independent observer and the mothers was low: 53% and 43%, respectively. Subsequently, the second mother self- recorded attention to inappropriate behavior. A 72% agreement was found for self-recording
attention to inappropriate behavior. The mothers' rates of attention to appropriate behavior increased 50% and 100%. The correlated increases in appropriate behavior for the children were a 29% increase to 80% and a 17% increase to a total of 90%. The 90% total was reached with self-recording inappropriate attention, following the first procedure. By self-recording an undesirable behavior, inappropriate attention, the desirable behavior was maintained above base line.

An increase in the frequency of a behavior self-recorded was obtained by Gottman and McFall (1972). Highly truant high school 10th graders were directed to self-record oral class participation or the urge to orally participate. Using a very interesting statistical analysis in a time-series design, a significant increase in self-recorded oral class participation was obtained. No similar increase was noted for the urge to participate. McKenzie and Rushall (1974) investigated the effectiveness of self-recording on attendance, late arrivals, early departures, work units and total laps completed for a swimming team. A behavioral analysis was utilized in two studies on 9-16 year old boys and girls. The results were as follows: absenteeism dropped by 45%, late arrivals decreased by 63%, early departures stopped, and work units plus total laps increased 27% on the average. A check on the accuracy of self-recording indicated a 100% correspondence between observer and swimmer.

In summary, a variety of behaviors across different populations seem to be amenable to influence by self-recording alone. It is of interest that the effect of self-recording is not directional, e.g.,
several behaviors increased when self-recorded, on-task (Broden et al., 1971), oral participation (Gottman & McFall, 1972), attention to appropriate behavior (Herbert & Baer, 1972), and attendance and performance (McKenzie & Rushall, 1974), and some self-recorded behaviors decreased, talk-outs (Broden et al., 1971; Lovitt, 1969), attention to inappropriate behavior (Herbert & Baer, 1972), hitting (Lovitt, 1969), and late arrivals and early departures (McKenzie & Rushall, 1974). Though some behaviors did not change, room cleanliness (Fixen et al., 1972) and the urge to talk (Gottman & McFall, 1972), the demand characteristics involved in self-recording may be an influencing factor.

The relative ease of implementing self-recording strategies is compelling, but the durability of the results is a question incompletely answered: Broden et al., and Gottman and McFall reported continued effects after two to three weeks; McKenzie and Rushall observed the resultant changes to be durable after 12 months. Of considerable interest is the accuracy of self-recording and the relationships between accuracy and behavior change. Herbert and Baer obtained as low as 43% correspondence with an observer and as high as 73%. Fixen et al., obtained an accuracy of correspondence at 50% before training and 79% after training. Neither study found any relationship between the accuracy of self-recording and behavior change. Other studies that combined self-recording with self-evaluation and self-reinforcement obtained a similar measure to the post-training student-observer correspondence: Bolstad and Johnson (1972) 71%; Glynn, Thomas, and Shee (1973) 76%; and Santogrossi, O'Leary, Romaczyk (1973), and Kaufman
and O'Leary (1973) correlations as low as -0.12. The reported results cogently question the a priori acceptance of the accuracy of self-recording, but self-recording seems to be a reasonably effective self-control procedure.

Self-recording may in part be an effective behavior change procedure due to the resultant increased awareness. Though learning with or without awareness is an unresolved issue, the importance of cognitive variables is not. Bandura (1969) has indicated that cognitive mediation may be more important than external variables in behavior change. Cognitive processes has long been thought to be central to self-control, but research has only recently focused on cognitive mediation.

Cognitive Mediation and Self-Control

Research focusing on cognitive variables that influence the performance of self-control has addressed several issues. The implications of the writings of Luria (1961) have been receiving much attention. Luria (1961) hypothesized that language influences behavior in three stages:

1. Children under two years of age are influenced somewhat by speech, i.e., the speech of others can initiate, direct and control, but not stop or inhibit ongoing behavior;

2. Later, the motor component of a child's speech helps to initiate behavior, but doesn't inhibit behavior;

3. Last, the semantic content becomes dominant, internalized, and directive.
Bern (1967) found some support for Luria's contentions. Three and four year old children were presented with a Luria type task: twelve lights were present on a box which had a clown's face decorated on it; a spring-loaded lever designed as the clown's tongue protruded. The task was to count, and press the appropriate number of times, dependent upon the presence or absence of lights as cues. Bern found that four year olds could perform the oral counting, but lever pressing was not facilitated. Luria (1961) concluded that the motor component facilitation emerged at four or five years of age. Bern implemented a training procedure for the three year olds and improved their skill to the point there were no differences between the three and four year olds. Bern inferred a learning deficit for the three year olds, rather than a developmental deficit.

Support for Luria's stage specific behavior was also found by Meichenbaum and Goodman (1969) with reflective and impulsive kindergartners and first graders. The task was pressing a footpedal when 12 blue lights were on, or not pressing when 12 yellow lights were on. The effectiveness of either covert or overt self-instructioning was assessed. The impulsive children were found to have less verbal control over motor behavior and used private speech in a less instrumental way than the reflective students. With covert instructions, impulsive children had significantly less control of inhibiting behavior. Forcing the first graders to voice the self-instructions reduced their performance, but enhanced kindergartner's performance. While these results seem to support Luria's conclusions, a study by
Miller, Shelton, and Flavell (1970) did not entirely provide support. The apparatus and task were Luria type: a box with a squeeze bulb and two lights, blue and yellow; the task was to squeeze the bulb when the appropriate light came on, don't squeeze when the other lights were on. The effect of "squeeze," "don't squeeze," and combined instructions were assessed. A sample of several ages was utilized: four groups of forty boys and girls: I, 3-0 to 3-4; II, 3-5 to 3-9, III, 3-10 to 4-3; IV, 4-8 to 5-4. Partial support for Luria's conclusions were found. The youngest children did better with overt self-instructions; but the control group, with no self-instructions, obtained the highest performance. Also, there was no interaction between age and conditions, i.e., the expected increased effectiveness of the semantic content of the self-instructions with older children did not obtain. The authors hypothesized that these results may have differed from Luria's because a repeated measures design was not used. Bem (1967) and Meichenbaum and Goodman (1969) did not use repeated measures, and they partially supported Luria's findings, so the results are conflicting.

Meichenbaum (1975) has described using an apparatus and task very similar to that of Miller et al., (1970), with the same negative results. The choice of a footpedal for the Meichenbaum and Goodman (1969) study was due to the Luria type apparatus breaking prior to the start of the experiment.

Meichenbaum (1975) suggested that the task may be an important variable. Luria (1961) described an apparatus with lights on a box
with a squeeze bulb and the task was repetitive squeezing or not squeezing. Studies using the squeeze bulb and the colored lights as discriminative stimuli have not found the tested mediational factors to obtain the predicted interaction with age and self-instruction: Bem (1967) and Meichenbaum and Goodman (1969) did not use a squeeze bulb, Miller et al. (1970) did.

The first two stages of Luria's Theory have received some support (Bem, 1967; Meichenbaum and Goodman, 1969; and Miller et al., 1970). The age specific behaviors described by stage three have received equivocal support from the same studies which supported stages one and two.

The implications of a much different theory were investigated by McKaughan (1975). Dulany's Theory of Propositional Control (1968) has obtained support with adults (Dulany, 1968), but McKaughan evaluated three different age levels of children: 4-1/2; 8-1/2; and 11-1/2. The highly formalized theory was expressed in a mathematical model which included the following variables:

(BH) - Behavioral hypothesis;
(RHd) - What response gets reinforced;
(RHs) - What is the correct significant response;
(BI) - The behavioral intention;
(RSv) - The subjective value of the reinforcement;
(MC) - Motivation to comply;
(R) - Performance;
(a and b) - Regression coefficients.
The values for the first six variables were obtained by asking six questions during the performance of the task. The task was selecting the correct rectangular card with the color dimension varied. For example, a question used was "Now think about these last cards again. Was there any special color that was right or wrong?" The answers to the questions indicated which values (+1.0, 0.0, or -1.0) were assigned to the subjects and predicates of the stated proposition. The value assigned depended on the extent the subject correlated with correct color, and the extent the predicate indicated affirmative action, e.g., when yellow is the correct color, the values would be as follows for these two statements: "Say (+1.0) yellow (+1.0)," "Don't say (-1.0) red (-1.0)." The value of a proposition is the algebraic product of its subject and predicate. The values were then used in the following equations to predict performance or relationships among the variables.

BH (Behavioral Hypothesis) = RHd (Distribution of Reinforcement) \times RHd (Hypothesis of Reinforcement Significance).

I (Behavioral Intention) = (RHd [Distribution of Reinforcement] \times RSv [Subjective Value of Reinforcement]) a (Regression Coefficient) + (BH [Behavioral Hypothesis] \times MC [Motivation to Comply]) b (Regression Coefficient). R (Performance) = BI (Behavioral Intention). The theoretical predictions were found to be highly correlated with the obtained performance of the 8-1/2 and 11-1/2 year olds, i.e., the children's scores and predicted results based on BH (behavioral
hypothesis), BI (behavioral intention), and R (performance) were significantly correlated. The results for the 4-1/2 year old were not adequately described. The results were interpreted by the author to be consistent with the body of evidence indicating that awareness is an antecedent to conditioning. The theory seems promising, yet little research has been generated by the theory.

Bem (1967) demonstrated support for a learning deficit for the three year olds in her sample. Several other studies have investigated the effectiveness of self-instructions on the performance of different behaviors with different populations. Palkes, Stewart, and Kahana (1968) found that training in the use of oral self-instructions improved the Porteus Maze performance of hyperactive boys. The boys read self-instructions aloud during the performance of different tasks: The Matching Familiar Figures Test; the Embedded Figures Test; and the Trail Marking Test. The training generalized to the Porteus Maze performance and reduced errors, but to establish whether voicing the instructions was necessary, another study followed. Pales, Stewart and Freedman (1971) used the same training tasks with hyperactive boys, but voiced self-instructions were compared with silently read self-instructions. The oral self-instructions were found to be significantly more effective and the silent self-instructions were as effective as a no self-instruction control group. The authors noted that the boys did not appear to verbalize during the task. Palkes et al. (1971), inferred that the most parsimonious explanation may be that recall was facilitated by voicing the self-instructions.
This inference seems congruent with the findings of Meichenbaum and Goodman (1971). The authors implemented two related investigations: the first assessed the effectiveness of a training procedure to reduce the deleterious effects of impulsivity on problem solving; the second evaluated the relative contribution of two of the training components. Reflective or impulsive eight year olds, from a class for children with poor self-instruction made up the sample. Three groups were used: cognitive self-guidance; attention control; and an assessment control. In the cognitive self-guidance group, the experimenter modeled self-instructions while performing a task. The child then self-instructed aloud with the experimenter's direction. Then, the child was asked to perform the task while whispering the self-instructions. Finally, the child was asked to covertly self-instruct while working. The self-instructions involved:

1. Questions about the nature and demands of the task;
2. Answers to these questions in the form of cognitive rehearsal and planning;
3. Self-instructions in the form of self-guidance while performing the task;
4. Ways of coping with errors and failure;
5. Self-praise.

The training tasks were increased in difficulty over the four, 30-minute training periods. The tasks included copying line patterns, coloring figures within boundaries, following sequential directions taken from the Stanford-Binet intelligence test, completing pictorial series as
on the Primary Mental Abilities Test, and solving conceptual tasks on the Raven Matrices test. The modeling and self-instruction treatment proved highly effective in reducing impulsiveness and decreasing errors on some of the measures. Using the same basic design, the authors evaluated the effectiveness of modeling self-instructions alone compared to modeling plus the step-wise self-instruction training as utilized in the first study. One training session was used. This training consisted of 20 minutes involving eight practice trials. The modeling group alternated between the experimenter's demonstrating the task and self-instructions, and the child's performing and imitating. The experimenter provided encouragement and praise for accurate modeling. A modeling plus self-instruction group received explicit self-instruction training with the oral self-instructions faded to covert self-instructions over the eight trials. The results indicated that both experimental groups reduced impulsiveness on the Matching Familiar Figures Test, but the modeling and self-instructions were effective in reducing errors, also. The results do seem to support Luria's conclusions about the important regulatory function of self-speech.

O'Leary (1968) also found that oral self-instructions were effective in regulating behavior. After first grade boys had been taught how to perform a discrimination task, the task was then presented so the children could earn tokens while working alone. Half of the children received no special instructions; the others were directed to self-instruct whether they could or could not depress a key on each trial, which self-reinforced with a token. The children were instructed to
depress the key only when the "right" one was present. The self-instructions were much more effective than no self-instructions: 94% of the boys who used self-instructions did not transgress.

Oral self-instructions with limited training have proved effective in increasing performance (Bem, 1961; O'Leary, 1968; Palkes et al., 1968, and Palkes et al., 1971), especially with younger children (Meichenbaum and Goodman, 1969; and Miller et al., 1970). More intense or extended training has proved effective in influencing performance with covert self-instruction (Meichenbaum and Goodman, 1971). A somewhat different method was assessed by Blackwood (1970). A mediation training procedure inferred from behavior theory was compared to a punishment procedure to reduce rule infractions by eighth and ninth graders. The children's disruptive behavior had been reduced, but not eliminated, by a token system. The mediation training was designed to increase self-control

... by conditioning verbal behaviors, such as descriptions of consequences of target behaviors, to mediate between temptation and target responses as self-produced discriminative stimuli and to follow target responses as conditioned reinforcers (p. 251).

The training procedure was the copying of an essay organized around four questions:

1. What did I do wrong;
2. Why is that behavior inappropriate;
3. What should I be doing;
4. What are reasons for the appropriate behavior.
The training followed six possible steps:

1. Copy the appropriate essay twice and return it by the deadline if no argument followed the assignment, only one essay was required;
2. If step one was not met, then two copies were required;
3. If step two was not met, then four copies;
4. If step three was not met, then detention;
5. If step four was not met, then two days' detention;
6. If step five was not met, then the principal was involved and severe punishment was recommended.

The training was administered in stages:

1. The first two times a child misbehaved, essays were assigned;
2. For the third through sixth infractions, the student was asked to paraphrase the essay, after school;
3. On the eighth through tenth infractions, the student was kept after school and required to write the essay in his own words from memory; the teacher provided prompts;
4. The final stage for continued misbehavior required the student to sit in his regular seat and orally describe the situations which instigated misbehavior and state how he might be tempted again; sometimes the student was directed to rehearse desired behavior and describe the consequences.

The punishment procedure involved an essay of equal length and difficulty as the mediation essay, but it described the workings of a steam engine. The punishment essay was assigned as was the mediation
essay. The results indicated that only the mediation training obtained significant gains and was more effective than the punishment procedures in reducing disruptive behavior. A similar study was undertaken to reduce disruptive lunchroom behavior by MacPherson, Benjamin and Hohman (1974). Three different procedures were compared in the reduction of disruptive lunchroom behavior in an elementary school: behavior modification techniques; a punishment essay and behavior modification; and a mediation essay and behavior modification. The essays, steps, and stages were basically the same as used by Blackwood (1970). The most effective treatment was the mediation essay and behavior modification next most effective were the behavior modification techniques alone, and the punishment essay and behavior modification were the least effective.

To this point, the described studies supporting relevance of self-instructions have generally followed from three different theories: Blackwood (1970), Dulany (1968), and Luria (1961). Other variables have received attention through systematic experimentation, more because of interest in the dependent variable than from explicit theoretical implication.

Resistance to temptation was the dependent variable in two studies which investigated the influence of the content of self-instructions and the relevance of who's voice uttered the instructions. Hartig and Kanfer (1973) compared the relative effectiveness of four different self-instructions with children three to seven years of age:

1. Verbalizing the positive consequences of a non-transgression;
2. Verbalizing negative consequences of a transgression;
3. Verbalizing instructions not to transgress;
4. Verbalizing nursery rhymes, i.e., irrelevant content;
5. No verbalizations.

The dependent variable was the time a child waited before turning around to look at toys, an act forbidden by the now absent experimenter. All three groups using relevant verbalizations were found to be equally effective, and significantly more effective, than either irrelevant self-instructions or no verbalizations. There were no differences between positive or negative directions, nor were there sex differences. Older children did make better use of the verbalizations, but they tended to prefer covert verbalizations. A follow-up study by Kanfer and Zich (1974) used the same dependent variable with children from 4-10 to 6-3. The variables assessed were whether the child heard his own voice saying tape recorded instructions or heard the experimenter's voice. Also evaluated was the effect of the experimenter's presence or absence during the latency period. The results indicated that the child's voice was more effective in sustaining the resistance to temptation than either the experimenter's voice or no voiced instructions. The absence of the experimenter resulted in longer latencies. Boys obtained longer latencies than girls; this was not found in the Hartig and Kanfer (1973) study. These two studies indicated that task relevant self-instructions combined with low external control, resulted in the most effective self-control.

Cognitive factors related to inhibition have also been investigated by Mischel. A series of studies beginning with Mischel and
Ebbesen (1970) systematically evaluated variables related to delay of gratification. Mischel and Ebbesen (1970) measured the effect of having an immediate reward, a more attractive delayed reward, or both in the child's sight. The children were 3-6 to 5-8 in age. The authors found that the conditions, with either one or two rewards present, were equal in effectiveness, but when no rewards were present the children were able to delay much longer. It seemed that not having the reward in view or in close proximity increased self-control. Mischel and Moore (1973a) did a similar study, but they compared the effectiveness of presenting slides of the reward to presenting slides of comparable rewards which were not to be received. The slides of the relevant rewards increased delay. The authors hypothesized that having a picture of the reward may increase information. The children might be saying to themselves "I know what I'll get," "I know what the contingency is." Having the reward in close proximity may increase frustration, e.g., I can't eat it," "I can't touch it or play with it," thus resulting in a response very difficult to inhibit. In order to evaluate whether the hypothesized frustration had a debilitating effect on self-control, Moore and Mischel (1973b) replicated their previous study with one difference, before the delay interval, the children were instructed to think about the consumatory qualities of the reward, i.e., the qualities involved with eating, touching or playing with the reward. The added sensitization to the consumatory qualities of the reward significantly reduced the children's delay of gratification. The most recent study, by Mischel and Baker (1975),
seems to more definitively assess the relevant ideation. Children 3 to 5 years of age were asked to choose between a pretzel now or several marshmallows in 20 minutes. The children could stop waiting whenever they chose, but while they waited they were instructed to play a "Think About" game. The children were asked to think about either consumatory qualities, i.e., taste, chewiness, sweetness, et cetera, or transformational qualities, i.e., color, shape, size, et cetera. The groups were also asked to think about either the reward they had chosen or the other reward. The variables then were consumatory versus transformational ideation and ideation regarding relevant versus irrelevant rewards. As predicted from the frustration hypothesis, consumatory ideation directed at the relevant reward was the least effective in increasing delay of gratification. Whether thinking about relevant or irrelevant rewards, focusing on transformational qualities facilitated delay most effectively.

These results were supported in a study by Ebbesen, Bowers, Phillips and Snyder (1975). Additional variables were found and the complexities of factors involved were made clear. Ebbesen et al., (1975) used the traditional "forbidden toy" paradigm. The child rated the attractiveness of several toys with the most attractive then forbidden for play. The child played with the less attractive toys. Then the experimenter obtained another toys rating and the child was then free to play with all toys. Several studies have found that under mild threat, the forbidden toy is devalued more than under severe threat (Aronson and Carlsmith, 1963; Freedman, 1965; and Turner and
Wright, 1965). Dissonance theory seems to best explain the phenomenon. However, the authors observed that a self-control interpretation was plausible. Following from studies like Mischel and Ebbesen (1970), what the child attends to may reflect the amount of frustration and resultant devaluation as a self-control strategy rather than dissonance reduction: mild threat may be strong enough to prevent play, but not ideation regarding the forbidden toy; severe threat may prevent both the play and ideation. The authors evaluated three levels of threat and two levels of toy attractiveness. They found that when the toys were all similar in attractiveness, the devaluation was directly related to severity of threat, the opposite of what was predicted by dissonance theory. The dissonance effect only obtained in the dissimilar toy condition. In a second experiment, two levels of threat were used and a comparison of either directing the child's attention to either the "liking" of the forbidden toy or to the child's lack of playing with the toy (self-perception). This provided a test of the frustration hypothesis (liking) or a self-perception hypothesis (recall of past behavior). When liking was emphasized, a dissonance effect obtained; however, the self-perception hypothesis was only partially supported, since a larger devaluation did not occur under high threat. So self-control of attention, to reduce frustration, received greater support. The results were not clear and the authors clearly stated the need for further research.

Following from a different line of research is an investigation of the preference for delayed rewards related to future time perspective.
Klineberg (1968) tested three hypotheses:

1. Compared to those who prefer immediate rewards, children who consistently select delayed rewards are more realistic and consistent in their outlooks on their personal future;

2. There is no relationship between a child's ability to choose a delayed reward during relatively short time frame and measures of the length of perspective on the distant future;

3. Children who consistently choose delayed rewards will show greater everyday preoccupation with future events and less with present occurrences, than children selecting immediate rewards.

Boys from Paris, France were sampled. One-half of the sample were children from schools for the behaviorally disordered, the others were from private boarding schools. The age range was 10-3 to 12-8. Four measures of future time perspective were derived using Thematic Apperception Cards, interviews, and scaling methods. Two measurements were made of delaying capacity. One, the child chose between immediate or delayed consequences in a story. Two, the child was delivered one of two choices: a small candy bar immediately or a candy bar twice as large one week later. The results indicated no difference in reward preferences between maladjusted and normals, nor among age levels. The results were as expected. Children who consistently selected delayed rewards were more consistent in their ordering of future events and showed less time variance, i.e., they were more realistic and moderate in their time estimates. There were no differences between the two
groups on any of the measures assessing length of future time perspective. The children who consistently selected delayed rewards showed more of a daily awareness of future events.

Much of the research dealing with cognitive variables related to self-control is directly or indirectly related to the writing of Luria (1961). Direct tests of theoretical implications were implemented by Bem (1967), Meichenbaum and Goodman (1969), Miller et al., (1970), and Meichenbaum (1975) and were generally supportive of the state related behaviors of stages one and two. However, the behavior seemed to be more a result of learning and maturation, since Bem (1967) was able to train three year olds to perform as only 4-1/2 to 4-1/2 year olds are expected, according to Luria (1961).

The learning deficit hypothesis inferred by Bem (1967) was supported as several investigators trained children to inhibit different behaviors across different populations. O'Leary (1968) found that first grade boys performed with less cheating when they orally verbalized the contingencies for self-reinforcement. Palkes et al., (1968), found that voicing self-instructions helped hyperactive fourth grade boys inhibit careless, impulsive errors; Meichenbaum and Goodman (1969) obtained similar results with a similar sample of kindergartners and first graders. Miller et al., (1970), using a wide range of ages in the sample, found only limited support for Luria's (1961) conclusions, since the semantic content of the self-instructions used did not facilitate performance. However, Meichenbaum (1975) described obtaining similar negative results and suggested the Miller
et al., (1970) findings may be an artifact of the apparatus and the task: none of the studies supporting Luria's stage related behavior used an apparatus and task similar to that of Miller et al., (1970).

Some general support for Meichenbaum's inference may be found in the studies dealing with ideation inhibiting gratification (Mischel and Ebbesen, 1970); Mischel and Moore, 1973a; 1973b; and Mischel and Baker, 1975). The above researchers were manipulating the content of children's ideation and measuring the influence on inhibition. Conceptually this seems very similar to Luria's (1961) use of self-instructions to inhibit motor responses. Granted, ideation, task, and dependent measures were different, but it seems difficult to understand how the delayed reward studies, with samples three to five years in age, two-thirds of which should not be able to benefit from the content of self-instructions (ideation), were able to consistently demonstrate that ideation very strongly influenced inhibition. Meichenbaum (1975) pointed out that the Lurai type task seems intuitively to allow for little opportunity to utilize the added mediational constructs: the task is squeezing or not squeezing, following the onset of stimulus lights. The wait for the delayed reward provides a highly appropriate task to benefit from appropriate mediation.

The lack of mediational concepts or learning deficit inferred by Bem (1967) seems congruent with the research by Blackwood (1970) and MacPherson (1974). Both provided a training procedure to provide mediational concepts which facilitated the inhibition of behavior, in this case, rule infractions. Functionally the mediation training
was a punishment procedure: the step-wise training followed each rule infraction and effectively decreased the frequency of the infraction.

Blackwood's (1970) analysis of how the mediational training facilitated self-control seemed to be as follows:

1. The content of an essay that is memorized becomes a discriminative stimulus for the avoidance of inappropriate behavior in a probable infraction situation and for the performance of appropriate behavior;
2. After successfully inhibiting the inappropriate behavior, the child would have matched the standard for acceptable performance and the successful matching would function as a self-reinforcer.

Due to the design, no conclusion can be reached regarding the function of the mediation essays content in Blackwood's (1970) study. Hartig and Kanfer (1973) found that it did not matter whether self-instructions emphasized what to avoid or what to do, inhibition was facilitated as long as the instructions were task relevant. Therefore, the relevance of the content may be a factor. There may be an additional variable operating. Part of the mediation essay provides a rationale for inhibiting behavior: question two, what should I be doing, and question four, why is the rule in effect. La Voie (1973) found that with junior high students, a parent's providing a rationale for inhibiting a behavior was as effective as an aversive noise, i.e., punishment. Therefore, the rationale may well be a variable adding to the punishment effect.
The most parsimonious explanation regarding mediational training may be as Palkes et al. (1971) observed regarding self-instructions, perhaps the experimental conditions simply facilitated recall. The training procedures implemented (Bem, 1967; Blackwood, 1970; Meichenbaum and Goodman 1971; and MacPherson et al., 1974), resulted in much more of an opportunity to memorize the appropriate behaviors and instructions. The other studies cited used instructions simple enough to be utilized immediately, e.g., Hartig and Kanfer (1973); instructions were also presented visually (Palkes et al., 1968; 1971) or by tape recorder (Kanfer and Zich, 1974).

Most of the studies cited presented simple tasks. Also, the self-instructions or influenced ideation did not require the generalization of self-control. The experimenter usually presented simple activities, which when later measured for performance changes, were the dependent variables. Four studies provided training on tasks which were similar, but different from the dependent variable; Palkes et al. (1969; 1971) and Meichenbaum and Goodman (1969; 1971). Both Palkes studies and Meichenbaum and Goodman (1969) obtained results which seemed to indicate that, if a child recalled what to do, it was sufficient to improve performance. However, Meichenbaum and Goodman (1971) cogently demonstrated that recalling what to do was necessary, but not sufficient for improving the performance of complex tasks. Unfortunately, Meichenbaum and Goodman (1971) used different dependent measures compared to their 1969 study. Therefore, the relationship between self-instructions, and task complexity and generalizability were not clarified.
Though recall was a factor, attention was shown to be very important in facilitating inhibition, e.g., awaiting a delayed reward. The most effective procedures were to have a child wait, after selecting the delayed reward, with no visual cues related to the selected reward (Mischel and Ebbesen, 1970). When slides of the delayed reward were viewed, inhibition was facilitated more than if slides of comparable unselected rewards were viewed (Mischel and Moore, 1973a). Neither slides of the relevant or irrelevant rewards were helpful, if the child was instructed to think about the consumatory qualities of the reward (Mischel and Moore, 1973b). Mischel's investigations have cogently supported the hypothesis that young children are much more effective in awaiting a selected delayed reward if they think about transformational attributes which may facilitate recall, as opposed to consumatory qualities, which may increase frustration (Mischel and Baker, 1975). A similar process was assessed by Ebbesen et al., (1975). The authors found that the attractiveness of the toys and the saliency of the child's liking or past behavior with the toy, interact in a complex manner with the level of threat to inhibit play with a forbidden toy. But what the child thinks about is crucial. Two studies on the resistance to temptation utilizing a different paradigm, have indicated that task relevant self-instructions are effective in strengthening the inhibition to not turn and look at toys (Hartig and Kanfer, 1973; Kanfer and Zich, 1974), but if the child is alone and hears his own recorded voice repeating the instructions not to look, the inhibition is strongest (Kanfer and Zich, 1974). The Kanfer and
Zich study was the only one to deal with the child's involvement in the task. Several studies have removed the experimenter during the measurement of the dependent variable to reduce the child's motivation to conform to demand characteristics (Mischel and Liebert, 1966; Stouwie, Hetherington and Parke, 1970; and Peskay and Masters, 1971), but Kanfer and Zich (1975) found that when the experimenter was present, the child apparently assumed less responsibility for conforming to the explicit instructions and was less inclined to resist looking at the toys. These three studies and much of the self-control research may have ignored a variable which influenced the main effects being studied.

Dulany's (1968) Theory of Propositional Control was utilized by McKaughan (1974) with a concept attainment task. The possibility of using different tasks and applying the theory exists, or for the testing of competing explanations, e.g., the O'Leary (1968) study used a concept attainment task and could possibly be analyzed by Dulany's theory. Also, Dulany's theory allows the investigation of variables investigated by other researchers, e.g., the effect of self-instruction, the value of different reinforcers, the effect of contingent versus non-contingent self-reinforcement and the magnitude of a child's behavioral intentions. To this point, little research has been stimulated by the theory, however.

A variable receiving little attention in self-control research is future time perspective (Kleinberg, 1968). In an interesting study, Kleinberg found that there was a significant relationship between children's perceived reality of personal future events and everyday preoccupation with the future. Several of the studies on preference
for delayed rewards used much younger children than Klineberg's 10-6 to 12-6 aged sample, e.g., 3-6 to 5-8 years of age (Mischel and Ebbesen, 1970), future time perspective may be a factor.

Much of the self-control research regarding cognitive mediation has addressed itself to improving performance or strengthening inhibition. Self-reinforcement has usually been considered an integral part of this self-control mediational process, but to this point, has received little attention per se.

**Laboratory Studies in Self-Control**

Non-applied settings have been the focus for much of the research with children on self-reinforcement, and, to a lesser extent, self-criticism. Two basic paradigms have been utilized in this research: modeling and directed learning. The modeling paradigm usually involves adult models exhibiting explicit criteria for self-reinforcement on a bowling game. The pins are not in view and an experimenter determined score is presented after each roll. After the child observes the model's self-praise and self-reinforcement, the child plays alone. The dependent measures usually involve the number of self-reinforcements utilizing the modeled criterion and the child's verbalizations. The directed learning paradigm usually involves the child receiving experimenter administered reinforcement during an activity requiring the child to respond. Then the child assumed the responsibility for self-reinforcing. The acquisition or transmission of patterns of self-reward has been investigated using modeling, while the effects
of variables on self-reinforcement have been explored by the use of the directed learning paradigm.

A great deal of research using both paradigms followed a classic study by Bandura and Kupers (1964). The authors were interested in how the imitation of self-reward criteria would be influenced by viewing models of different age and sex. Children 7 to 9 years of age viewed an adult or a peer model. The models used either a high or low criterion for self-praise and self-reinforcement (M & M's) i.e., scores of 20 or above (high) or scores of 15 or below (low). The possible scores were 5, 10, 15, 20 and 30. The low criterion children self-reinforced twice as often as the high criterion observers. Though no sex of model and sex of subject interaction was found, the adult models were more effective in transmitting self-reward criteria and verbalizations. None of the control children expressed positive or negative self-evaluative statements, but a significant number of the experimental group reproduced precisely the self-approving or self-critical statements made by the model. This study cogently demonstrated that children can learn self-reinforcement patterns vicariously. Though it was difficult for the subjects to evaluate their performance, due to the variability of the scores and the lack of normative information, the question was raised, what would happen when the model and subject's competence was divergent.

The same authors (Bandura and Whalen, 1966) expanded a replication to include success or failure experience, three competency levels of models' compared to subjects' performance (superior, equal or inferior),
and three different self-reward criteria: high criteria by the superior mode; moderate by the equal; and low by the inferior. Children 8 to 11 years old had free access to candy during the bowling game. Children observing a superior model rejected the high criterion and generously self-rewarded using low criteria. This rejection of the modeled standards was seen as congruent with a social-comparison theory. According to the theory (Festinger, 1965), people tend to select models who are similar to their ability and reject those who are dissimilar. The children observing inferior models self-reinforced significantly more often for low scores. Children who experienced failure tended to self-reinforce less, but the success-failure variable was unrelated to the amount of self-reinforcement. No matter what the modeling condition or success-failure condition, all children in the high performance condition self-reinforced more than the other conditions. The only sex difference seemed to be that boys used more verbal self-reward.

Mischel and Liebert (1966) also found that social learning influenced subsequent self-reward. Using the bowling game, fourth graders were exposed to two procedures: a modeled criterion was followed by urging and praise from the model for the child's use of a specific self-reward criterion; then the child demonstrated the game to a same sex second grader. Half the sample played the game alone prior to demonstrating to a peer, the others afterward. The modeled criteria and direct training were as follows:

1. High criterion modeled—the child guided to use a low criterion.
2. Low criterion modeled—the child guided to use a high criterion.

3. High criterion modeled—the child guided to use a high criterion.

The discrepant modeling-directed learning sequences parallel the often quoted maxim "Do as I say, not as I do." The results indicated that the children were most stringent in self-rewarding in the consistent condition, i.e., modeled high-guided high. Otherwise, the children utilized the guided criterion, i.e., when the two criteria were discrepant the children did as told. These results generally held when the child was either a performer or a demonstrator. Thus, the results seemed to indicate the children did as told rather than as they observed. These results were partially supported by Rosenhan, Frederick and Burrowes (1968) with children the same age. The authors used a similar design to that of Mischel and Liebert (1966), but the subject-peer modeling aspect was excluded. An adult played the bowling game and then directed the child to use a standard consistent or inconsistent with the modeled standard for self-reward. An important difference between the two studies was the attractiveness of the reward: Mischel and Liebert (1966) informed the subjects that they would earn tokens exchangeable for "valuable prizes." Rosenhan et al., (1968) explicitly described contingencies which resulted in earning gift certificates ranging in value from five cents to 50 cents. The experimental conditions were designed so that the more attractive rewards were obtainable only by cheating. As in Mischel and Liebert's study, stringent-stringent
obtained the highest use of the stringent standard, but 40% of the subjects still violated the standard. If the children only used the stringent criterion they could only earn a 10 cent certificate. The fewest norm violations were in conditions where the child was told to use lower standards thus allowing the sanctioned earning of a 50 cent certificate. The highest rate of cheating occurred when the child was urged to use a stringent standard and the model used a lenient criterion. It does appear that children tend to do as told rather than as they observe, when divergent standards are used. It is also apparent that reward magnitude is an important variable.

These results were supported by McMains and Liebert (1968) using a different design to investigate the effects of divergent criteria between two different adult models. Fourth graders were exposed to a model who stated a stringent standard for self-reward, then played the bowling game and used either a stringent or a lenient criterion. A second model then played using one of the criteria, resulting in the following combinations:

1. Stringent-stringent-stringent;
2. Stringent-lenient-stringent;
3. Stringent-lenient-stringent;
4. Stringent-lenient-lenient.

The child played the bowling game alone, after viewing the models. The results indicated the combinations were most effective in transmitting the stringent standard in the above order: stringent-stringent-stringent was the most effective with each combination less effective than the one
above it. The consistent stringent condition had the highest proportion of stringent self-rewarders, but generally, it appeared that the most recent standard was adopted. To better evaluate whether concordance (consistency) or recency described the more effective sequence, Hildebrandt, Feldman, Solomon and Ditrichs (1973) added a different person to give the rule. The authors reasoned that the first model's statement of the rule reduced the equivalence of the two models. This might explain why the consistent and more recent stringent-lenient-lenient sequence was less effective than the stringent-lenient-stringent sequence. Hildebrandt et al., (1973) had the experimenter explain the bowling game to the child and state either a lenient or a stringent standard. After the child played a game alone, a model played a game, then the child played a game alone before and after a second model. There were eight possible combinations. Generally, concordance better described the most effective sequence for teaching a stringent criterion. The most effective was stringent-stringent-stringent; the least effective was lenient-lenient-lenient. Though concordance seems a better predictor of learning, the phenomenon appears incompletely understood at present.

Conflicting adult models, conflicting adult and peer models, and relative age and competence seem to be important in the transmission of self-reward patterns. To further explore adult-peer modeling conflict and model characteristics, Bandura, Grusec, and Menlove (1967) evaluated the influence of the model's nurturance. The authors hypothesized the peer model's influence would be negated by the increased
attractiveness of the nurturant adult model. Half the subjects interacted with the model who was approving and rewarding prior to the modeling experience, the others interacted with a model behaving non-nurturantly. All children were exposed to an adult model who performed the bowling task at a consistently superior level and adopted a high criterion of self-reward. Half the models in each nurturance level received praise for adopting a stringent criterion, the others did not. Half the children in each subgroup observed a stringent adult and a peer using a low criterion for self-reward, in order to evaluate the effects of divergent models. The trend of adopting a low standard when divergent standards are presented was supported. But contrary to expectations, the non-nurturant models obtained the most stringent self-reward. Apparently, nurturance was interpreted as permissiveness by the children. The models receiving praise and the lack of a conflicting peer modeling experience also increased the effectiveness of the stringent criteria being adopted.

The social power of the model was investigated by Mischel and Liebert (1967). Power was defined as having the ability to bring about consequences affecting the child. The model with social power was introduced to the second and third graders as the vice president and general manager of the toy company who made the bowling game. The child was told there were extra games and there was a good chance some of the children would get one free. The rewards were candy and nuts which were non-contingent on bowling scores and no relationship was implied between bowling scores and possibly receiving an extra bowling
game. The powerful model told half the sample that a particularly good chance existed for that child to get a game. Then the model alternated bowling trials with child, displaying lenient standards and directing the child to use a stringent standard. The model left and the child played alone; then the experimenter returned and negated the incentive, i.e., told the child there was no chance of getting a bowling game. The child then played alone again. The children viewing the powerful model were more stringent before negation, though after negation both groups were equal. The authors pointed out that negation was uncontrolled. Adding another power group and not negating the incentive would give clearer results.

A study similar to Bandura and Kupers (1964) by Bee and Colle (1967) compared direct learning versus modeling of either high or low standards for self-reward. The elementary school subjects obtained significant learning effects except in the high standard modeling group. Unexpectedly, the level of self-reward in the bowling game was equal between the high standard modeling group and the control group. The authors hypothesized that either competence or socio-economic status (SES) may have influenced the results. These variables were controlled in a second study (Colle and Bee, 1968). Competence for both model and subject was determined by programmed scores, e.g., high competence was the child's receiving two scores less than 10 and seven scores greater than or equal to 20. The models praised themselves and took candies for scores greater than or equal to 20, but self-derogated and refused candies for less than 20. The results did not support Bandura
and Whalen (1966). No significant relationship was found between model and subject competence; however, low SES subjects rejected the high standards modeled and the high SES children accepted them. Thinking that perhaps the low SES subjects were less intelligent and had not yet learned the standard, a correlation between standard setting and IQ was obtained, but found to be not significant.

Both studies by Bee and Colle (1967; 1968) included self-praise or self-criticism by the model, e.g., "That was a good score," however, the contingency was not stated, e.g., "That was a good score, that deserves a candy." Liebert and Allen (1967) varied the level of rule structure, i.e., the explicitness of the contingency, and the reward magnitude, along with comparing direct training versus modeling and sex differences. The results indicated that reward magnitude did not result in differences in immoral behavior, i.e., when playing the bowling game alone, the children exhibited no differences in rule violations related to reward magnitude. However, the more explicit the rule, the more stringent the self-reward criterion. Also, the more rule structure present, the more verbal self-reinforcement and self-criticism was imitated. There were no significant effects due to sex, and unlike Bee and Colle (1967), none due to modes of training. Liebert and Allen (1967) used two levels of rule structure. A finer discrimination was sought by Liebert, Hanratty, and Hill (1969); they used three levels: high structure: social approval and a statement of deservingness; low structure: statement of rule for self-reward; and moderate structure: social approval only. Effects due to sex
differences and training modes were also evaluated. No differences were found between boys and girls nor for modeling compared to direct training, but the more structure, the greater the subjects' adherence to stringent self-reward criteria. The moderate rule structure was equally effective as the low structure (the moderate structure condition was a replication of the Bee and Colle [1967] study), which did not support previous results, i.e., the moderate structure in direct training was not superior to the moderate structure modeling group, as found by Bee and Colle (1967).

Bandura and Kupers (1964) found that self-reward patterns could be transmitted vicariously. Though Bee and Colle (1967) demonstrated that with moderate rule structure, direct training was more effective in teaching self-reward standards, Liebert and his colleagues did not find support for the findings (Liebert and Allen, 1967; and Liebert, Hanratty, and Hill, 1969). Liebert's studies found that direct training was equally effective as modeling, and no sex differences were found. While investigating different variables, Bandura and Perloff found that boys seemed to be more receptive to experimenter-administered reinforcement than self-reinforcement, so sex differences may be a factor among children 7 to 10 years of age. This was not supported by Liebert and Ora (1968), but children 8 to 10 years of age were found to learn self-reward standards used on the bowling game equally well as a result of direct or vicarious training. Unlike Liebert and Allen (1967), the authors found significantly more generous self-rewarding with high incentives compared to less attractive rewards.
The previously described studies which compared the influence of modeling experiences which were consistent or inconsistent with a stated rule (e.g., Hildebrandt et al., 1973 and McMains and Liebert, 1968) might be considered a comparison of direct training (rule statement) and modeling. These studies have found that an interaction effect occurs, but generally the direct training aspect predominates. However, the research comparing the effectiveness of direct training versus modeling has found both equally effective (e.g., Liebert et al., 1969 and Liebert and Ora, 1968).

The model's characteristics have been found to have very strong effects on the transmission of self-reward and self-criticism. Several studies previously cited included model characteristics as independent variables. Generally the results seemed to indicate that adults were more effective models than peers (Bandura and Kupers, 1964), the sex of the model was not a factor (Bandura and Kupers, 1964), low competent children rejected highly competent models (Bandura and Whalen, 1966), highly competent models were rejected by low SES children (Colle and Bee, 1968), powerful models were more effective (Allen and Liebert, 1969), and nurturant models were less effective (Bandura et al., 1967). The above studies dealt with self-reward; some studies dealing with other dependent variables also investigated model attributes.

Preceding the Bandura et al., (1967) study on nurturance was a study by Grusec (1966) on self-criticism. This study will be discussed in more detail later, but Grusec found that significantly more
self-critical behavior was learned when the children observed a nurturant model. These data were not supported by Bandura et al. (1967); therefore, self-criticism and self-reward may well be the result of different learning processes. Allen and Liebert (1969) varied the amount of information children received on the model's alleged experience on the bowling game. No main effect for experience was found, but experience seemed to facilitate learning, e.g., the group with the highest level of learning observed an experienced model. The other variable studies was incentive. The effects of incentive proved very powerful and the authors inferred that experience effects may have been masked. Stouwie et al. (1970) included sex of model in an investigation of the effect of achievement orientations on self-reward and self-criticism. High and low achievement oriented third and fourth graders observed male or female adults playing the bowling game, then the model directed the child's playing, and then the child played alone. The model self-rewarded for scores of 15 and 20 with self-praise and self-criticism for scores below 15. The model directed the child to self-reward only for a 20 score. The female model obtained the most stringent self-rewarding. These results do not support Bandura and Kupers (1964), but the designs are different. Since these are the only studies dealing directly with the sex of the model, the need for replication is evident.

Two studies by Thelen and Fryrear (1971a; and 1971b) investigated the effect of the model's race. Black and white institutionalized delinquents 15 to 17 years of age viewed a video tape of a black or
white male model attempting a pursuit rotor task. The two experiments were the same but one involved boys (1971a), the other girls (1971b). The reward was chips worth five cents, exchangeable for edibles. The model used either a liberal (self-reward after every trial) or a stringent (self-reward on half the trial) standard. The white liberal model was imitated more than the black liberal model by boys, but both liberal models were equally imitated by the girls. Both boys and girls were equally influenced by the stringent models irrespective of race. A post hoc comparison found the black boys to be significantly more influenced by the liberal white model compared to the black girls. The authors inferred that attributed power, sex roles of institutional staff or missing parents may be factors. Unlike the previously cited studies involving second through sixth grade public school children, the studies by Thelen indicate a need for research regarding different ages and different populations and the transmission of self-reward standards.

Self-criticism and self-punishment have received relatively little attention in the last ten years. Some of the previously cited studies measured self-criticism as a dependent variable and found that modeling was effective in transmitting the behavior (Bandura and Whalen, 1966; Bandura and Kupers, 1964; Liebert and Allen, 1967; Liebert et al., 1969; and Stouwie et al., 1970), but no differences were reported for self-criticism as opposed to self-rewards or self-praise nor were any published studies found dealing only with this topic.
Mischel and Grusec (1966) found that nurturance and power influenced the rehearsal of aversive and neutral behaviors differentially. Nursery school children were introduced to a powerful (new) teacher or a low power (substitute) teacher. After a nurturant or non-nurturant experience with the teacher, the children observed the teacher model neutral behaviors. Then while the child was playing with a toy which broke, the teacher labeled the child a "store wrecker" and punished the child. The child was then left alone and imitated behavior was recorded. Neutral and aversive behavior was significantly more prominent in the high power-high nurturance condition. The highly nurturant model obtained more neutral behavior imitated and the high social power model seemed related to the imitation of aversive behavior, e.g., no aversive behavior was imitated in the low social power conditions. Previously, Mischel and Liebert (1967) found that social power increased model effectiveness in transmitting stringent self-reward criteria, but Bandura et al., (1967) found that highly nurturant models were less effective in modeling stringent self-reward standards. It appears that power enhanced the imitation of self-reward and self-criticism and nurturance enhanced the imitation of neutral behavior, but not stringent self-reward. Grusec (1966) also evaluated the effect of model nurturance on the adoption of self-critical behavior. Kindergarten children were exposed to a nurturant or non-nurturant model, then the model and the child played a game with toy soldiers and a toy donkey: the child pushed the unseen donkey out of the soldiers' way with a rod; mistakes, controlled
by the experimenter, were signalled by a buzzer. Whenever the child erred, the model punished with either withdrawal of material reinforcers (WOMR), i.e., tokens, or withdrawal of love (WOL), e.g., "You're a hurter, I'm not happy with the way you're playing," and the model bowed her head and looked unhappy. The reward (tokens or praise) was reinstated contingently upon self-criticism or noncontingently: both of these conditions included prompts for self-criticism. After 15 trials a generalization measure was obtained; then the next eight trials were an extinction measure. Unlike Mischel and Grusec (1966), exposure to nurturant models resulted in significantly more self-critical (I'm a hurter") behavior. WOL interacted with the other variables in a complex manner. The high nurturant models using WOL and reinstating rewards contingently were the most effective, and the low nurturant model using WOL and noncontingently reinstating rewards was the least effective. The model using WOMR was of intermediate effectiveness. The children whose rewards reinstatement was contingent upon self-criticism showed the most generalization.

The same relative effectiveness of the procedures continued through the extinction trials with no diminished responding.

The influence of subject variables on self-criticism were tested by Herbert, Gelfand, and Hartman (1969). High or low self-esteem fourth graders were assigned to modeling or no modeling groups. The model displayed self-critical remarks and self-punishment (giving up tokens) for low scores on the bowling game. The children exposed to models showed significantly more self-punishment, but there was no difference in self-criticism. Interestingly, there were no significant
main effects for self-esteem. This seems contrary to expectations, since children with low self-esteem are considered prone toward self-
derogation. Using an ambiguous visual discrimination task, Haynes and Kanfer (1971) found that when training only involved externally administered criticism, third and fourth graders of low-academic rank were subsequently less self-critical than high ranking students. In a similar second part of the experiment, when externally-administered praise, criticism or no feedback were compared, only the low ranking children used more self-praise or self-criticism than the no feedback control. Apparently, low achievers are more prone to use self-criticism and self-praise, while high achievers tend to use self-criticism only. The only other studies using self-criticism exclusive of self-reward were Mischel and Grusec (1966). Both studies combined modeling and direct training, but in fact, the aversive behavior was not modeled in either study. Herbert et al., (1969) used only modeling and did not obtain learning. Thus, drawing into question the efficacy of the transmission of self-critical behavior via modeling. A study by Thelen (1970) is relevant. Thelen had elementary school children observe an adult model make self-blame statements after failing a card sorting task. Three groups then undertook the same task, each receiving different feedback:

1. Positive consequences: "Oh, you're doing o.k."
2. Negative consequences: "Oh, you're not doing it right"
3. No consequences.

Seven months later, the task was presented again. The negative
consequences and no consequences groups showed more self-blame than
the positive consequences or a no model control. Therefore, it seems
that self-criticism can be transmitted through modeling and can be
instigated after an appreciable period. Inferences regarding self-
criticism are difficult to make due to the paucity of research.

Subject variables have received some attention in the previously
cited studies. Stouwie et al., (1970) found that achievement moti-
vation was an important variable for girls, not boys, i.e., girls with
high achievement motivation deviated less from a proscribed high
standard for self-reward, even when the model had been self-indulgent
while playing the bowling game. Herbert et al., (1969) found self-
esteeem, as measured by pencil paper instruments, was not a factor in
the transmission of self-critical behavior, but class rank was (Haynes
and Knafer, 1971). Other studies have found that the subject's com-
petence (Bandura and Whalen, 1966), SES (Colle and Bee, 1968), and
race (Thelen and Fryrear, 1971a; 1971b) influence the transmission
of self-administered consequences.

Incentive level or reward magnitude studies have not supported
Liebert and Allen (1967) who found no difference between high and low
magnitudes of reward: children were given tokens with no explanations
(low magnitude) or they were told the tokens would be exchanged for
a valuable prize (high magnitude). Liebert and Ora (1968) argued
that the difference between high and low incentive was slight, so
they replicated the incentive manipulation, but the children working
for prizes viewed their potential reward prior to the bowling game.
High incentives resulted in significantly more self-reward standard violations compared to low-incentives in direct training or modeling. Rosenhan and Burrowes (1968) obtained similar results when direct training preceded successive modeling. Allen and Liebert (1969) used modeling and also found high incentives, i.e., having the child view the potential prizes before bowling, resulted in more deviations from the modeled standard.

One modeling study bears reporting since self-reward was studied in association with another dependent variable, performance. Kunce and Thelen (1972) utilized Thelen's paradigm: black and white delinquent adolescent boys viewed liberal or stringent self-rewarding black or white models on a pursuit rotor task. Actual time on target was measured and found to be higher after viewing stringent models. There was no relationship between subject self-reward and performance. Viewing the stringent model seemed to increase motivation.

The studies reported have included modeling either alone, in direct comparison with direct training or combined with direct training. The studies to follow deal with self-reward in non-modeling paradigms. With few exceptions, the research involves subjects' receiving experimenter-determined reinforcement followed by the subjects' self-determined reward.

Social comparison processes were demonstrated to be operating in some modeling studies (Bandura and Whalen, 1966; and Colle and Bee, 1968). Masters published several studies which investigated the influence of social comparison processes on several variables related
to directed learning and self-reward. Equity theory was the context of three studies to be described (Masters, 1968). Equity theory, an extension of social comparison theory (Festinger, 1957), proposed that individuals compared themselves to others in terms of what each individual contributes in a situation and what is derived. Thus a comparison is made of the ratio of input/to output, which holds for each individual. For example, when a worker has the same wage but more seniority as another, inequity results. Inequity then promotes behavior changes which either alter the ratios, obtaining equity, or else terminate the comparison. Masters hypothesized that children receiving less or more reinforcement than a partner would attempt to regain equity when given an opportunity, i.e., those children receiving less would generously self-reward, those receiving more would self-reward less, and those receiving equal reward as a partner would self-reward equal to the non-comparison control. The sample consisted of children 3-10 to 5-2 in age. Same sex pairs, partners were a year younger than subjects, underwent one of three experimental conditions:

1. Low: subject received 9 tokens--partner received 54--reservoir none;
2. Equal: both received 9 tokens--reservoir 45;
3. High: subject received 9--partner received 3--reservoir received 51.

The experimental task was a game (Paymaster) where the child was asked a series of simple questions that he was sure to know. After the Paymaster game, the subject was asked to do simple pencil-paper mazes.
Then the child was given 3 minutes and told to take as many rewards as wanted, i.e., non contingently. The rewards present were pennies, tokens, and squares of paper. A second study replicated the first, but replaced the mazes with a second opportunity to play Paymaster and the subjects dispensed rewards to the experimenter and to themselves. No partner was present. A third study replicated most of the first study replacing the mazes with playing Paymaster again with the experimenter asking the question, but the child dispensing the rewards to the partner and himself. The hypotheses were only partially supported. No significant differences were obtained for the different reinforcers or for any of the conditions in experiment three. Generally the predictions held for boys, i.e., those in the low comparison group, later self-rewarded more than the equal or high groups. However, the low boys self-rewarded more than their control. The girls, however, self-rewarded more than the equal group in both the low and high groups.

To clarify the above results, Masters ran a second study (Masters, 1969) investigating, in addition, the influence of varying reinforcer magnitudes. The controls were altered slightly also. As before, a non-social comparison group was used. The child received 9 tokens, and the reservoir receiving 54, but instead of having the child receive 9 and none go to the reservoir, in a second control, the child was told 54 were going for the other nursery school children so they could get a nice prize too. This was in effect a social comparison group. After the Paymaster game, the partner had returned to
the nursery, the child was asked to give the experimenter the number of tokens each of the following was worth: a penny; a slinky; and crayons. The tokens were returned after each turn. The child and the experimenter played Paymaster again with the child distributing 4 tokens as desired, after each question. The results for girls were the same as in the 1968 study. Generally, children in the low and new control group, i.e., the children who received 9 and the "others" received 54, self-rewarded significantly more than the high, equal or non-social comparison groups. These three groups were not significantly different. There was no correlation between self-reward and the value of the reward. However, children in the low conditions were willing to trade significantly more tokens for a penny than those in the social comparison or high conditions. Two interesting findings were the consistent sex effect, girls self-reward more after both the low and high equity manipulations as in Masters, 1968, and the high level of self-reward in the new social comparison group. This seemed to indicate that a partner did not have to be present for the social comparison effect to obtain.

Masters (1971) used the same design, but used male and female experimenters and evaluated the effect of social comparison on giving tokens away (altruism) to a competitor or a friend. Five year olds had four year old partners during the equity manipulation as in Masters, 1969. Then the child played "Give Away," i.e., how many tokens the child would donate to the partner, with the partner present, and then how many would be donated to a friend, a partner not present.
The child then played Paymaster with the experimenter as a partner; the child was given 4 tokens to distribute after each game question. Previously (Masters, 1968), no social comparison effect was found for sharing with a peer present, so none was predicted here. The previous results were not replicated as the children in the low equity conditions gave significantly fewer tokens to a partner. Also, the Master, 1969 results were replicated showing that an absent partner ("other children in the nursery") has a powerful social comparison effect. The social comparison condition resulted in less altruism to a friend. The previous sex effect for girls (Masters, 1968; 1969) was not replicated as no sex of subject effects were found. But sex of experimenter was highly significant: children more generously self-rewarded in the female experimenter's presence and donated fewer tokens to a friend in the presence of a male experimenter. Little relationship was found between self-reward and altruism, but much support for their both being influenced by social comparison processes was established.

Masters (1973) also compared the contingent and non-contingent self-reward of preschool and second grade children. The reinforcer value was also assessed. The child and younger partner were administered the Matching Familiar Figures Test in such a fashion that response correctness was ambiguous. Children were rewarded contingently or non-contingently, during the standard equity manipulations. The child then worked mazes without a partner and self-rewarded. Last the child estimated how many tokens a penny was worth. The older children
were not influenced by the equity manipulations. A powerful age effect was noted upon self-reward and the value of the tokens; only younger children self-rewarded more in the social comparison and valued the tokens less. Young children in the low and high social comparison conditions self-rewarded more than the equal condition. There were no significant differences found for contingent or non-contingent self-reward nor was there any main sex effect. Thus it appears that by the time children have reached the second grade they are not as sensitive to the comparison with a generalized other and do not seek subsequent increased self-gratification.

Modeling has been found to be effective in transmitting self-reward standards (Bandura and Whalen, 1966; Bandura and Kupers, 1964), direct training and modeling have been found to be equally effective in transmitting self-reward standards (Liebert and Allen, 1967; Liebert and Ora, 1968; and Liebert et al., 1969). The first direct test of direct training or experimenter-administered reinforcement and self-reinforcement on performance was implemented by Bandura and Perloff (1967). Children 7 to 10 years of age were assigned to one of four groups:

1. Self-selection of a criterion for self-reward;
2. Experimenter selected criteria, i.e., these children were yoked to those in group one for self-reward standard;
3. Rewards were given non-contingently before the task was begun; these children were yoked to group one for amount of reinforcement;
4. No reward.

The children in the self-reward group used an apparatus that had a crank on its side which when turned eight times resulted in a light coming on. Four lights were labeled 5, 10, 15, or 20. By moving a selection switch, the child could select any of the four numbers (criteria). After the appropriate number of turns of the crank were completed, a chime sounded, the light came on, a button was pushed by the child and a token was dispensed into a bowl at the front of the apparatus. For example, if the child selected 15, after 24 turns the third light (the 15 light) would slide into the bowl. After the child was explained the task and took a practice trial, the experimenter left and the child selected a standard and began. Each child was allowed to change the standard once. Externally administered reinforcement was equally effective as self-reinforcement and both obtained significantly more wheel turning than the control groups. Surprisingly, the children did not maximize their reward/work ratio. Boys did significantly more cranking in the externally rewarded conditions. Since the incentive control group "inherited" their tokens in a lump sum before the cranking was performed, the timing was procedurally incorrect. In order to better assess non-contingent self-reward as a control, Liebert, Spiegler, and Hall (1970), rewarded the children at randomly selected standards. The apparatus was similar to that of Bandura and Perloff, but two revolutions rather than eight, separated the different self-reward standards: 5 points = two revolutions; 10 points, four revolutions; 15 points, six revolutions;
and 20 points = 8 revolutions. To make the task more realistic a time limit (4 minutes) was put on the task. The same age sample was used, but another control was added. The no reinforcement group in the Bandura and Perloff study did hear a chime, so a no chime, no reinforcement group was added. Two levels of reinforcer values were also assessed: children in the high incentive group viewed objects to be earned worth up to $2.45; the others viewed objects worth up to 10 cents. The experimenter explained and demonstrated the task, then the child was alone for 4 minutes. An unlimited number of standard changes was allowed. The results indicated the higher the reward value the higher the performance. The high incentive group was significantly higher than the control groups. The no reinforcement, no-chime control was significantly higher than the chime, no reinforcement group. However, these controls were not different from the no incentive or low incentive group.

Incentive level did seem to influence performance level. The self-determined group and the externally-rewarded groups were equal in productivity, so the Liebert et al., study essentially supported Bandura and Perloff (1967). There is some question, however, whether the no reward control group would be significantly different from the experimental groups used by Bandura and Perloff because of the relatively low incentive value of the tokens. The sex differences found by Bandura and Perloff were not replicated, but two differences between the studies may be factors:

1. Bandura and Perloff did not control the amount of reinforcer for the externally-administered group, the reward
standard was controlled, thus a fixed ratio schedule was in effect; Liebert et al., used a fixed interval schedule with the amount of tokens controlled, not the standard;

2. Different sex experimenters were used.

The effect of self-reinforcement on extinction and generalization were studied by Johnson (1970). First and second grade boys with poor attention skills were trained on a match-to-sample task.

The children were assessed on the task while music and attractive toys were available for distraction. Those who were inattentive were subjected to the following sequence of activities:

1. Continuous experimenter-reinforcement for attention on task;
2. One-third of the sample were (a) taught to self-reinforce (b) not reinforced or (c) were continued with experimenter-rewards;
3. Continued performance with either self-reinforcement, experimenter-reinforcement, or no reinforcement;
4. Extinction;
5. Experimenter- or self-reinforcement re-implemented;
6. Generalization in the classroom: cross out 5's in rows of random numbers.

The reinforcements were points, later exchanged for toys or candy.

The self-reinforcement group was initially more resistant to extinction, but not significantly so. Both experimenter and self-reinforcement groups were more effective than a no reinforcement control, but experimenter-and self-reinforcement were not significantly different. There was no significant generalization to the classroom. Weiner and
Dubanoski (1975) asked second, third and fourth grade boys and girls to select a schedule of reinforcement (fixed ratios of 1, 2, or 4 responses for each token), then measured the effect on subsequent extinction. The task was dropping a hard rubber ball into a hole in the top of a box. The ball exited the box at the bottom after a three second delay.

After twenty responses, extinction began. The self-selectors were yoked to children of same sex and grade. The dependent variables were time to extinction and the number of responses during extinction. Children in the self-selected schedules condition were more resistant to extinction and responded more than the externally-selected group. The different schedules demonstrated a partial reinforcement effect i.e., a fixed ratio of 2 was more resistant than a fixed ratio of one, and a fixed ratio of 4 responses per reinforcement was the most resistant to extinction. An interaction between sex and schedule was found. Boys using the fixed-ratio 4 schedule responded more and were more resistant to extinction than girls. These results seem similar to those of Bandura and Perloff (1967) and male experimenters were employed in both studies.

Only one published study has purported to investigate the reinforcing effect of self-reward (Montgomery and Parton, 1970). A "guessing game" was explained to children 8 to 11 years of age. The apparatus had a transluscent screen with three buttons and a cup below, and a lever to the side. On each trial the screen was illuminated with one of five colors. The child was told to "guess," and press the button which
went with each color, then pull the lever if an apparent correct match had been made. Half the sample received a penny in the cup after each lever pull. The children were told they could not keep the pennies. The children underwent 85 trials. The results indicated a significant increase in the probability of the same response to two consecutive presentations of the same color when the match was self-reinforced. There was a simultaneous decline in probability of the same button push to the same two stimuli, with no self-reward. Also, the lever alone was significantly less rewarding than the lever and penny condition. Apparently, learning did occur with self-reward.

Masters had demonstrated that age is related to reward value (Masters, 1969; 1973) and social comparison influences (Master, 1973). Kanfer (1966) studied the influence of age upon the frequency of self-reward. He reasoned that not only age, but level of incentive and class standing would influence the number of self-rewards in a situation where only the subject apparently knew the correctness of the response. Children from the second through eighth grade were asked to guess a number between 0-100; the experimenter then selected a number lottery-fashion and announced the number. The child then announced whether that had been the self-selected number. The probability of a correct choice was .01, so all correct matches were considered undeserved self-rewards. Points or candy were the rewards. Younger children self-rewarded significantly more, as did children ranked in the lower half of their class. Children who viewed an adult model self-reward on 3 of 5 trials, also self-rewarded more than those
not viewing the model. Some of the third and fourth graders were subjected to a second experiment. The children partially learned a visual discrimination task. After each trial, the child pressed a button self-rewarding a point or candy. The ambiguous task was performed equally well by high or low ranked children, but the low ranked took significantly more undeserved self-rewards. There was no difference based on incentive level. A significant relationship was found for the children's taking undeserved self-rewards in both experiments. Age, class standing and incentive level seem very much related to the level of self-reward.

Kanfer and Duerfeldt (1968) partially replicated the Kanfer (1966) study and added two variables. The authors reasoned that the child's commitment to a guess would influence cheating and also the public labeling of undeserved self-reward would influence cheating. The variables were as follows:

1. Labeling cheating versus no labeling;
2. Writing the guess versus no writing;
3. High versus low class standing;
4. Four grade levels; 2nd - 5th grades;
5. Five trials.

The task was guessing a number between 0-100. As in Kanfer's (1966) study, younger children and low ranking children self-rewarded more. Requiring the writing of the guess (increasing commitment) reduced cheating, but labeling undeserved self-reward as cheating had no effect. Haynes and Kanfer (1971) investigated the effect of different
types of feedback on the self-reward behavior of children with high or low class standing. The child was trained to make a visual discrimination on an ambiguous task with the experimenter only labeling "bad guesses." Then the child was instructed to press one of two buttons after each choice: one button should be pressed on a "Very Good Guess" and the other on a "Bad Guess." The high achievers emitted significantly more self-critical remarks during the task than the low achievers. In order to evaluate the subsequent influence positive feedback would have, another experiment with the same children followed. Three types of feedback were delivered on the same task with new stimuli: positive, negative and no feedback. Only those children with low class ranks obtained differential results. Compared to the no feedback control, more self-criticism was found in both the positive and negative feedback conditions for the low ranking children. The high ranking children declined more in self-criticism, but there were no significant differences among the three conditions. Self-reward seems very much influenced by age, class standing and commitment with younger children and low class rank children self-rewarding more. Self-criticism seems to be more an attribute of the high achiever, though low achievers readily learn to use negative self-evaluation statements.

One way to interpret the influence of age and class rank on self-reward is that younger children and low ranking children self-reward to compensate for previously low levels of reward. This is consistent with equity theory. A previously discussed study by Bandura and Whalen
(1966) subjected their subjects to success or failure experiences. Generally, they found that children who experienced failure, self-rewarded less though only the inferior model elicited significantly low self-rewarding. This was interpreted then as the successful "self-congratulated" themselves. However, the control group experiencing failure, self-rewarded higher than two-thirds of the success groups. Apparent "self-therapeutic" self-rewarding was also operation. Masters (1975) has reasoned that an effective control for self-reward is non-contingent self-administered reward, thereby reducing normative and evaluative factors. Mischel, Coates, and Raskoff (1968), implemented such a control. Third and fourth graders played the bowling game. Though the scores were identical for all, half the sample did not get as many "good" scores as the others. The children then worked on easy mazes and were told to take tokens whenever they wished (non-contingently). The children who experienced success self-rewarded more than the failure group, and those who self-rewarded more were more persistent on the mazes. Though there were no main effects for sex, boys were more persistent than girls after success. A "self-congratulatory" effect was evident, but no "self-therapy." The authors felt "goodwill" toward the experimenter may have influenced the results, so another experiment followed. Self-rewarding was made more private and the two tasks were presented sequentially, as before, or concurrently. Though the children were informed the bowling and mazes could be performed concurrently, all followed the same sequence as the sequentially directed group, bowling then mazes. Still, the
children increased self-reward after success, only in the sequential group. This time girls were more persistent after success; the opposite from the first experiment. Masters and Mokros (1974) have suggested that perhaps the children didn't see the self-reward as non-contingent, since the more self-reward, the higher the persistence, i.e., the reward may have been interpreted to be for persistence. Though no self-therapy was found at all, perhaps the self-congratulatory rewarding was lacking in the concurrent condition due to the two tasks not being seen as different. A study by Masters (1972) dealt with these issues. Contingent rewarding was compared to non-contingent, and the task used in the success-failure manipulation was again evaluated for half the sample; a different task was used for the others: a pursuit rotor was used for training and mazes were used as the different task. Children experiencing success, self-rewarded more than those experiencing failure. There was no difference between contingent or non-contingent in success. The failed children self-rewarded more with non-contingent self-reward and later on a different task, regardless of the contingency. Non-contingent self-reward occurred with higher frequency than contingent self-reward. Apparently, both self-congratulatory and self-therapeutic self-reward occur. Perhaps the children in the Mischel et al., (1968) study did not see the mazes and bowling as different tasks in the concurrent condition. The Masters, 1972 results seem to support Masters and Mokros (1974) suggestion that Mischel's non-contingent self-reward may have been seen as contingent.

Masters replicated his 1972 study and added two levels of SES and also sampled black children (Masters and Peskay, 1972). The results of
Masters, 1972 were supported for the white children, but not for the black. First of all, the low SES blacks and whites self-rewarded more than the high SES children, but there were no differences between low or high SES blacks. The blacks did not self-reward more than the whites but a trend toward significance was present (p > .10). The expected higher non-contingent self-reward did obtain for whites, but not for blacks. The results on SES and race support earlier results from a similar study by Peskay and Masters (1971).

Success and failure may well produce the type of self-evaluation often associated with a more enduring self-evaluation: self-esteem. If this is true, high self-esteem would be expected to obtain high self-reward and low self-esteem would result in low contingent self-reward on tasks associated with the esteem measure. Herbert et al., (1969) found no relationship between self-esteem measured on pencil paper instruments and the imitation of self-punishment or self-criticism. The relationship between self-reward and self-esteem in 7th graders was investigated by Reschely and Mittman (1973). After completing a pencil paper self-esteem inventory, the children were presented with ambiguous tasks. They were instructed to self-reward points (0-10) on the basis of their performance. Significant differences were obtained in the direction predicted by the success-failure studies, i.e., the higher the self-esteem, the more ambiguous the task, the lower the rate of self-reward. The self-esteem and self-reward relationship might be considered congruent with the idea that past success or failure associated with a task is related to present
self-evaluation and subsequent self-reward. Positive self-evaluations and success probably would be accompanied by positive or happy feelings, as negative self-evaluations and failure would be accompanied by sad feelings. The relationship between these feelings and self-reward were studied by Underwood, Moore, and Rosenhan (1973). A female experimenter instructed the children and a male experimenter induced the affect. The children were told that for one minute they would be able to take all the pennies they wanted from a bowl; 500 new pennies were in the bowl. Before they took the pennies one group was asked to think happy thoughts for 30 seconds, the other sad thoughts; a control group waited while the experimenter slowly counted to 30. Unlike the success-failure studies (e.g., Masters, 1972; or Mischel et al., 1968) or the expected self-esteem-self reward relationship (Reschely and Mittman, 1973) both positive and negative affect resulted in higher levels of self-reward than the control. A sex difference was found: girls self-rewarded higher following either affect induction; boys self-rewarded higher in the positive affect condition only. These data did not clarify previous findings, though they support the contention that affect along with success and failure is an important antecedent event related to self-reward. The relationships between race, SES, self-esteem, success-failure, affect and self-reward, and contingent on non-contingent self-reward, remain to be more completely studied.

A different, but very important antecedent experience was studied by Karoly and Kanfer (1974). Every statement of a contingency is an
implicit contract, so a child's prior history of contracts might be important. The authors studied the relationship between prior contractual experiences and self-control. Self-control was defined here as tolerance for noxious stimulation. Girls, 7-11 to 11-11, were asked to play scarecrow. The game required that they hold their arms out "perfectly straight for as long as possible." Next a contract was made for the performance of a monotonous letter-crossing task. Five trials with each contract were involved:

1. Kept contract: promised two candies per trial, paid two;
2. Broken contract—negative: promised three, paid two;
3. Broken contract—positive: promised one, paid two;
4. Double message: promised two, paid two and gave mild verbal criticism for performance;
5. Comparison group—played the game without contract or reward.

Next, the second experimenter conducted an interview assessing the liking, fairness, and honesty of several stimuli including the first experimenter. The first experimenter returned, played another game of scarecrow and then asked the child to choose between an immediate reward or a larger delayed reward. No main effects were significant. Group three, the broken contract—positive increased in self-control over baseline and the comparison group declined. The others remained stable. The authors suggested that the "warm glow" of success may have aroused the children to increased self-control, but the other experimental groups didn't decline when expected. More research is needed on this topic.
The laboratory studies dealing with self-control have attended to self-reinforcement largely ignoring self-control per se, self-punishment, and to a lesser extent self-criticism. The acquisition of standards for self-reward has been demonstrated with modeling (Bandura and Whalen, 1966; and Bandura and Kupers, 1964), and modeling has been found to be as effective as direct training in transmitting self-reward criteria (Liebert and Allen, 1967; Liebert and Ora, 1968; and Liebert et al., 1969). When direct training and modeling are combined the resultant effect is powerful, except when the direct training and modeling are inconsistent (e.g., McMains and Liebert, 1968; and Mischel and Liebert, 1966). When the stated and modeled self-reward criteria are divergent, the concordance of the criteria rather than recency is a better predictor of which will be adopted (Hildebrant et al., 1973). The explicitness of the self-reward contingency is also important (Liebert and Allen, 1967; and Liebert et al., 1969). Social comparison processes have been found to influence the imitation of self-reward mechanisms (Bandura and Whalen, 1966). There are several apparent attributes of the model that enter into the social comparison, i.e., competence (Allen and Liebert, 1969; Bandura and Whalen, 1966), SES (Colle and Bee, 1968; Masters and Peskay, 1972; Peskay and Masters, 1971), race (Thelen and Fryrear, 1971a; 1971b), power (Mischel and Liebert, 1967), nurturance (Bandura et al., 1967), sex (Masters, 1971; and Stouwie et al., 1970), and age (Bandura and Kupers, 1964). Whether direct training or modeling is utilized, several subject variables have been identified: age (Kanfer, 1966; Kanfer and Duerfield, 1968);
Masters, 1973; self-esteem (Reschely and Mittman, 1973); class rank (Kanfer, 1966; Kanfer and Deerfield, 1968); achievement motivation (Masters and Peskay, 1972; Stouwie et al., 1970; Thelen and Fryrear, 1971a; 1971b); affect (Underwood and Rosenhan, 1973), and measures related to sex of subject: responsiveness to externally-selected reward criteria (Bandura and Perloff, 1967); sensitivity to social comparison (Masters, 1968; 1969; 1973); persistence (Mischel et al., 1968); achievement motivation (Stouwie et al., 1973); rejection of modeled standards (Thelen and Fryrear, 1971a; 1971b); affect (Underwood and Rosenhan, 1973); and performance and extinction (Wiener and Dubanoski, 1975).

Self-criticism and self-punishment have been found to be influenced by variables that affect self-reward, but not always in the same way. Bandura et al., (1967) found that nurturant models were relatively ineffective in transmitting stringent self-reward patterns. Grusec (1966) obtained the highest performance of self-criticism when nurturant experimenters rewarded self-criticism. Power seemed to increase the imitation of self-criticism while nurturance had no effect (Mischel and Grusec, 1966). Herbert et al., (1969) demonstrated that self-punishment could be acquired through modeling, but that self-esteem was not related to the acquisition of either self-punishment or self-criticism. High achievers seem to make more use of self-criticism than low achievers, but low achievers appear to be more easily influenced to increase or maintain self-criticism (Haynes and Kanfer, 1971).
Social comparison has been demonstrated to be an important variable in direct training (Masters, 1968; 1969; and 1973), but it does not seem to be a factor for children seven years and older (Masters, 1973).

Laboratory studies have not attended a great deal to the relative effectiveness of self- versus externally-determined consequences. Self-selected self-reward criteria seem to result in performance levels equal to criteria externally-determined (Bandura and Perloff, 1967; Liebert et al., 1970) and self-selected schedules of self-reward have been demonstrated to be more resistant to extinction than externally-determined schedules (Weiner and Dubanowsky, 1975). Increases in task relevant attention have been maintained equally well with self- or externally-administered self-reward (Johnson, 1970). Laboratory studies have shown consistently that self-determined consequences are equal to or superior to the externally-determined, but the number of published studies is small.

Several other variables seem to influence self-determined consequences. Success experiences precede increased self-rewarding compared to the self-reward performance of failed children (Bandura and Whalen, 1966; Masters, 1972; Masters and Peskay, 1972; and Mischel et al., 1968) and failure increases subsequent non-contingent self-rewarding or contingent self-reward for a dissimilar task (Masters, 1972). Self-criticism seems to be influenced more, by success or failure, with children ranking low in class standing (Haynes and Kanfer, 1971).
The value of reinforcers seems to increase with age and highly valued rewards are preferred by high SES children (Masters, 1973), but the research on the influence of incentive level upon performance, has obtained varied results. The self-rewarding following social comparison manipulations has not been demonstrated to be influenced by the value of the reward (Masters, 1968; 1969). The non-contingent self-reward on a maze game was also unaffected by the reward value (Peskay and Masters, 1971). However, performance has been shown to be directly related to incentive level (Liebert et al., 1970) and violations of modeled self-reward criteria are also directly related to incentive level (Allen and Liebert, 1969; Liebert and Ora, 1968). Though Bandura and Whalen (1966) found increased modeling with higher self-reward magnitudes, this has not been supported (Liebert and Allen, 1967; Masters, 1972). The conflicting results regarding reinforcer magnitude and value are unresolved and should benefit from continued study.

Other variables related to self-control receiving less attention are task ambiguity (Reschely and Mittman, 1973), commitment (Kanfer and Duerfeldt, 1968), and prior contractual experiences (Karoly and Kanfer, 1974).

Modeling studies have utilized contingent self-reward, albeit task ambiguity was often high (e.g., Bandura and Whalen, 1966). The directed learning studies have often used non-contingent self-reward, thus reducing normative influences, as does an ambiguous task. The relative effects of contingent versus non-contingent reward are therefore important, if the results from the different paradigms are
to be compared. When the social comparison manipulations were implemented with either contingent or non-contingent experimenter-determined self-reward, there were no significant differences between the two (Masters, 1973). No significant differences obtained, when the subsequent performances of self-determined contingent self-reward criteria and externally-determined non-contingent self-reward, were compared (Liebert et al., 1970). With both a maze game and a pursuit rotor, non-contingent self-reward occurred at significantly higher levels than contingent self-reward. So to the extent that the self-reward is contingent, lower levels of self-reward would be expected, compared to non-contingent self-reward.

The laboratory research seems to have stimulated a proliferation of variables. Because the variables are incompletely studied, they presently defy coherent organization. Social comparison theory has provided some structure, but the control and rigor characteristic of laboratory research has been uneven and unproductive in clarifying hypotheses or in delineating cogent mediational processes. Several areas are much in need of further study, e.g., self-punishment, self-criticism, reward value, incentive level, prior contractual experiences, commitment, and sex differences. The lack of control for sex of model or sex of experimenter is common. This may be in part a tribute to the classic research by Bandura. The Bandura and Kupers (1964) study found no differences due to sex of model and that apparently went unreplicated until Stouwie et al., (1970) where differences were found. It seems that the laboratory research is on the verge of finding
unifying principles for self-control phenomena, but already established is the efficacy of transmitting and maintaining self-control behavior. Now it is up to the applied researchers to demonstrate that the significant laboratory findings are meaningful in the natural environment.

Self-Control in Applied Settings

The application of self-control procedures has been evaluated extensively in educational settings. Though the relative efficacy of external- versus self-reward has received little attention in the laboratory, several applied studies have compared the two reward systems.

An applied behavioral analysis was implemented to evaluate the relative effectiveness of teacher-administered points or student-administered points (Lovitt and Curtiss, 1969). A 12 year old learning disabled boy from an Experimental Educational Unit, who had been on behavior modification programs for two years, was the subject. After a baseline measurement, the student entered into a contract with the teacher which specified nine tasks for which points could be earned. While the teacher designated contingencies were in effect, the teacher awarded points and recorded those earned. Next the student orally stated the self-awarded points in each of the nine areas and self-recorded the points. Then the teacher-determined contingencies were again implemented. The correct responses per minute were as follows: baseline 1.8; teacher-contingency 1.65; self-determined contingency 2.5; and teacher-contingency 1.9. With the apparent success and observed higher rate of correct responses per minute in the
self-determined phase, the program was replicated without a baseline. When the child was in the self-determined contingencies phase the second time, the author noted an apparent change. Though the results were the same as the first study for correct responses per minute, the amount of self-reward seemed higher compared to the teacher-determined phases. A third experiment was implemented to control for the amount of reward. The teacher first imposed contingencies as before with the same amount of reward as the first phase in experiment two. Then the teacher continued to determine the contingencies, but used the child's liberal criterion for reward in the self-determined phase of experiment two. Last, the teacher continued to determine the contingencies, but reverted to the more stringent reward criteria. The correct responses per minute were: Phase One 1.5; Phase Two (liberal reward) 1.2; and Phase Three 1.4. For this student, self-imposed contingencies resulted in increased academic rate and the amount of self-reward was unrelated to response rate. The studies by Masters (1968; 1969; 1973) seem congruent with the conclusion regarding magnitude of self-reward. Lovitt (1973) also evaluated the correct responses per minute subsequent to self-selecting assignment schedules or self-selecting different subject. Children from an Experimental Educational Unit were again utilized. A twelve year old boy self-scheduled morning activities. After a six week baseline the teacher limited the alternatives, but the child scheduled morning and afternoon assignment for the next five weeks. The morning assignment had to be finished before lunch and the afternoon assignment completed before the
child could leave school. During the following three weeks, the child sequenced the entire day's activities and leaving school was contingent upon the student's work being completed. Correcting his own work was the next task for the child. Next the child was shown his daily performance scores and he was soon asked to verbalize his relative performance. Viewing performance charts, charting in two subjects, charting in six subjects, and setting his own contingencies were also steps gradually phased in over the 29 week study. Except for the steps involving viewing performance charts and plotting performance in two school subjects, a steady increase in academic rate was obtained: Baseline 1.8 to 2.7; Step nine 3.7 to 4.2. No attempt was made to analyze the separate steps. The validation of the entire procedure was the goal. A second experiment involved two boys, 8 and 12 years of age. The first child had three 30 minute periods for evaluation: the teacher assigned reading or math alternately during the first period. The child was allowed to select reading or math the second period. In the third period, the teacher assigned the alternate subject to that presented in the first. The other child had three periods also: a 20 minute session in which the child selected reading or math; a 40 minute period where the teacher directed the child to continue with the same subject; and a 60 minute third period for the alternate subject to the first session. Three measures were taken each day. Two for the teacher-selected subjects and one for the child-selected subject. On 17 of 25 days the first child's performance was higher when he selected the subject; the other child's performance was
higher for the self-selected subjects 24 of 26 days. Thus, for these two boys the self-selection of the subject, appeared to increase academic rate. A third experiment assessed the effect of an eight year old boy's earning self-scheduling. During a baseline of correct responses per minute and errors per minute, while teacher imposed schedules were in effect, the correct rate decreased and the error rate increased. During the second phase, self-scheduling five activities was earned by the student by committing fewer than four errors during the phonic activities. The correct rate increased and the error rate decreased. For this child the contingent self-scheduling was effective. A fourth experiment combined several previously assessed steps. Nine children in the second grade were taught how to time, correct, and chart correct and incorrect response rates of programmed reading activities. After the training, an increase in correct rate and a decrease in error rate was obtained. The self-management procedures were continued when the children left the Experimental Education Unit and returned to regular class. The group performance continued to improve. The third step established access back to the Experimental Education Unit with a contingency. The academic performance increased again. The four experiments implemented by Lovitt (1973) demonstrated that self-scheduling, self-selecting subject, and charting can be motivating. The experiments did not include as much control for the experimental conditions as could have been provided, because the purpose was to demonstrate their efficacy.

A much more controlled study was implemented by Glynn (1970) which compared self-determined reinforcement to externally-determined
reinforcement in the classroom. Two other reinforcement procedures were controls: a group yoked to the self-reinforcement group for amount of reinforcement, and a no reinforcement group. The subjects were well motivated students with no behavior problems: 128 ninth grade girls. The dependent measure was performance on a 20 question, multiple choice reading test over history and geography readings. During a 10 day baseline, the students daily read a 500 word passage and then were administered the test. The children scored their own tests and recorded the number correct. Four more phases of 10 days each followed:

1) A first token phase included the experimenter-determined contingency; students received one token per four correct answers; the self-determined contingency group self-awarded from 0-5 tokens per test; the students yoked to the self-determined group; and a group that received no reinforcement throughout the experiment;

2) All reinforcement stopped;

3) Phase four, all three token groups self-determined their contingencies;

4) The last phase included a five day baseline and a five day reimplementation of the tokens;

5) A review test made up on items from the daily tests terminated the experimental procedure: seven tokens were available for self-reward on the review test.
The tokens were exchangeable for a range of prizes twice during the study: at the end of the first token phase and at the end of the study. The results indicated that self-determined contingencies were equally as effective as experimenter-determined contingencies. The non-reinforced children were consistently inferior to the experimental groups. The yoked control was as effective as the experimental group in the first withdrawal period. The authors attributed a reduced difference among reinforcement groups after the first token period to a reduced attractiveness of prizes exchangeable for tokens. After the first exchange of tokens, new prizes may have increased the value of the tokens. There is some question regarding any increased performance with higher valued tokens as Masters (1968; 1969; 1973) found no support for this inference. Performance-token ratios were also computed, i.e., the number of tokens self-awarded per correct response. A significantly higher number of correct responses per token were required by the self-determined group. There were no significant differences among the three reinforcement groups in amount of reinforcement. The yoked control group generally performed below the non-reinforced group. Glynn inferred the performance decrement of the yoked control group to be a result of the inconsistent reward experience in the first token period. The subsequent reward experiences were self-determined. Apparently, the unpredictable reward magnitudes reduced any incentive effect of subsequent self-determined contingencies.

Instead of using group comparisons, Bolstad and Johnson (1972) used applied behavior analysis to compare the effects of external-reward
versus self-reward on target behaviors per minute, i.e., talking out, physical aggression, and out of seat. The authors wanted to assess whether self-evaluation and self-monitoring would increase the resistance to extinction of self-reward over experimenter-reward. Neither had been demonstrated previously to be superior. On the basis of six days of observation, the four most disruptive children in 10 different combination first-second grade classes were initially selected for the study. Due to absenteeism and decreases in disruptive behavior the number was reduced to 33; a five subject control was later added. Each class was observed during the experimental manipulation, i.e., 30 minutes per day, for eight weeks. The baseline was Phase 1; Phase 2 involved three of four groups receiving external-reward of points according to the following contingency:

1) Fewer than five disruptive behaviors, 8 points;
2) Fewer than 10 disruptive behaviors, 4 points;
3) More than 10 disruptive behaviors, 0 points;

One group was not reinforced. Points were redeemable for experimenter controlled rewards: pencils, erasers, notepads, et cetera. Points were dispensed immediately after the experimental period and the prizes selected immediately, though the prizes were dispensed at the end of the day. In Phase 3, two groups received training in self-regulation:

1) The children received cards to self-monitor disruptive behaviors;
2) At the end of each session, the subject's record was compared to the observer's; if the tally was within 3
of the observers', points were awarded according to the contingencies in Phase 2; if the tally was not that close, two points were subtracted.

In Phase 4 the activities in all four groups continued, but the accuracy check between subjects and observer was discontinued. In Phase 5 all reinforcement was discontinued and one of the self-reward groups continued to self-evaluate and self-monitor. An additional control group of five students in one classroom, who all received no reinforcement, was added after the study had begun. No strong support was found for self-regulation being more resistant to extinction than external regulation. A significant difference in resistance to extinction between the two procedures was found, but difference scores were used. The authors cautioned that these results be considered in view of the difficulties with gain scores. The external- and self-regulation groups were equally effective in reducing the frequency of the target behaviors and they were consistently lower than the control groups. The self-regulation groups were consistently more effective in reducing disruptive behaviors and maintaining the low levels, but not significantly so. An interesting finding, was the difference between control groups. The control group with reinforced subjects in the same room was significantly lower in target behaviors than the control with no experimental subjects in proximity. There may have been a modeling effect. There was a measure of the discrepancy between the points deserved and the points self-awarded. An average of one unearned point per session was self-awarded, i.e., the deserved
average was 6.4, the average self-awarded was 7.4. The authors reported seven of the nine experimental subjects consistently over-rewarded themselves, the others accurately self-rewarded. Overall the accuracy of the self-observation was in agreement with the adult observer—71%.

The two point punishment implemented by Bolstad and Johnson for inaccurate self-evaluation during Phase 3, is often called response cost. Kaufman and O'Leary (1972) evaluated the relative effectiveness of response cost and reward procedures with a different population: adolescents residing in a psychiatric hospital. Two groups of 8 children were matched on age, sex, psychiatric diagnosis, IQ, reading grade, teacher liking, and initial level of disruptive behavior. Every pupil was observed during the 45 minute period for at least 15 minutes recording nine categories of disruptive behavior. One class followed the other after school. The reward or cost procedures were assigned by chance. Eleven teacher behaviors were observed daily also. The procedure was as follows:

1) Baseline, Phase 1, for 10 sessions;
2) Phase 2, tokens available according to reward or cost procedures, 9 sessions;
3) Phase 3, token withdrawal in both classes, 7 sessions;
4) Phase 4, tokens reinstated, as in Phase 2, 10 sessions;
5) Phase 5, class meeting times reversed to evaluate time of day effect 6 sessions;
6) Phase 6, self-evaluation: the subjects self-evaluated their behavior and rewarded or imposed response cost
themselves: reward class, 6 sessions; response cost, 7 sessions.

The reward procedure was as follows: the teacher announced the tokens earned (up to 10) and the rules to be followed by all students, at the end of each 15 minute segment of the 45 minute period. In the cost procedure, subjects began each 15 minutes with 10 tokens and the teacher announced the number lost and the specific infractions at the end of each 15 minutes. At that time, the actual tokens were put in or removed from containers at the front of the room and the teacher also recorded the numbers. The tokens were exchanged later in the day. The ratings and token manipulations were based on the teacher's observations. The results of the study found no significant differences between reward or cost procedures in the amount of work increase, reading achievement gain, or the decrease in disruptive behaviors, but these all improved, significantly. If the response cost had a disruptive effect, as might be expected, then increased disruptive behavior would follow the teacher ratings and subsequent token removal. The results indicated both procedures obtained increased disruptive behavior following the ratings and token manipulations, so the expectation was not met. The self-evaluation procedure was effective in maintaining low levels of disruptive behavior after self-reward was discontinued. This seems very encouraging as there was no training in self-evaluation and self-reward, per se. In fact, the authors reported a non-significant correlation between student and teacher ratings during this phase. Few pupils gave themselves lower than perfect
ratings, but some did. Even though this self-reward appeared to be non-contingent upon rule-following, the disruptive behavior remained low.

Several questions are raised by this apparent non-relationship between observed behavior and student self-reward. In order to better evaluate the effects of self-evaluation on disruptive behavior, and also to lengthen the time self-evaluation is in effect relative to the Kaufman and O'Leary (1972) study, Santogrossi, O'Leary, Romantzyk, and Kaufman (1973) implemented a study with the similar population. An after-school remedial reading class of 30 minutes duration met, with nine adolescent residents of a psychiatric hospital in attendance. Behavior ratings and token awards were made after each 15 minute block, i.e., twice per period. The same nine disruptive behaviors utilized in the Kaufman and O'Leary (1972) study were used by Santogrossi, et al., (1973). The 51 days of the experiment went through the following phases:

1) Baseline, 9 classes;
2) Self-evaluation by the students, 6 classes;
3) Teacher-determined tokens, token exchanged for prizes, 8 days;
4) Teacher awarded tokens or token penalties based upon teacher-student rating match, 1 day;
5) Teacher-determined tokens, tokens exchanged for prizes, 14 days;
6) Self-determined tokens, tokens exchanged for prizes, 5 days.
Phase 5 was implemented because of an increase in disruptive behavior at the end of Phase 4: the students apparently realized that they could self-reward non-contingently.

The self-evaluation alone, Phase 2, was ineffective in reducing disruptive behaviors. The teacher-determined token system was very effective, Phase 3. This low level of disruptive behavior was maintained for four days during the beginning of the self-evaluation and self-reward. The next four days described an increase to baseline, apparently when the subjects realized they could lie. Phase 5, the matching condition elicited anger, the highest level of disruptive behavior, and was quickly terminated. The subsequent reinstitution of the teacher- and self-determined tokens were much less effective. The authors reported some students actively non-cooperating, and though the behavior rates were substantially below baseline, they were higher than the first teacher-determined system. Teacher-student ratings correlated highly through Phase 3. After that, the students consistently overrated their behavior and the correlations were weaker. Given the length of the experimental manipulation and the population, the authors inferred that the self-determined reward was an ineffective procedure. The Kaufman and O'Leary (1972) study implemented a teacher-determined system for 25 days, Santogrossi et al., implemented a similar system for 9 days and were unable to replicate the length of maintenance with self-evaluation procedures. Therefore, the maximum period low disruptive behavior rates can be maintained with self-evaluation procedures is still an open question. Santogrossi et al.,
did demonstrate that a relatively short teacher-determined token system was not adequate in preparing adolescent psychiatric residents for self-control.

A very similar population was used by Drabman, Spitalnik and O'Leary (1973) to also evaluate the efficacy of self-evaluation to maintain low rates of disruptive behavior. Drabman et al., observed that Kaufman and O'Leary (1972) and Santogrossi et al., (1973) made the transition from teacher-determined reinforcement to student-determined reinforcement, abruptly. Drabman et al., designed a procedure which involved training and fading of external-control, so the transition was gradual.

The subjects were eight, 9-10 year old boys in public school classes for students with academic and emotional problems, who were enrolled in an after school remedial reading class. The 60 minute class was divided into 15 minute periods. One of the periods was selected at random to exclude the awarding of tokens, i.e., the period was a token control. Each child was observed during one randomly selected period. The disruptive behaviors measured were the same as in Kaufman and O'Leary (1972) and Santogrossi et al., (1973). The primary dependent measures were the number of disruptive behaviors per 20 second interval, daily progress in Sullivan readers and pre-post scores of the Reading Vocabulary scores of the California Achievement Test. The procedure involved eight phases:

1) Baseline, 11 days;
2) Teacher-determined reinforcement, 5 days;
3) Matching, a one point bonus for student, teacher rating match, 10 days;

4) Fading 1: four children's ratings were evaluated, 7 days;

5) Fading 2: two children's ratings evaluated, 7 days;

6) Fading 3: one child's rating evaluated, 4 days;

7) Fading 4: one-half the time one child's rating evaluated, 2 days;

8) Self-evaluation: self-reward of tokens with no penalties for over-evaluation, 6 days.

Whenever a child or group was evaluated for accuracy of self-rating, the children evaluated were selected at random. This check was presented as an honor. The tokens were exchanged daily for edibles or pennies.

The results of the study indicated that the self-evaluation procedure was effective in maintaining low levels of disruptive behavior. The baseline rate of disruptive behavior was .86, the token reinforcement reduced the rate to .28. A slight rise to .34 was obtained during matching, but by the end of the fading procedures the rate was .11. During the self-evaluation phase a rise to .19 was obtained which was still considerably lower than any of the first three phases. An unexpected result was a corresponding decrease in rate of disruptive behavior during the control period. In fact, the lowest rate obtained during the self-evaluation phase in the control period, .09. The authors inferred that the shortness of the control period and the juxtaposition with reward periods may have been responsible. This is
he first reported generalization of self-evaluation procedures, published. Also, the 30 days training period, i.e., the matching phase and four fading phases, seem to be the basic procedural difference between Drabman et al., and the other studies that failed to maintain low rates of disruptive behavior with self-evaluation (Kaufman and O'Leary, 1972; Santogrossi et al., 1973). Drabman et al., also did not specify the behaviors to be rewarded. The authors inferred that the vagueness may have elicited overcompensation from the subject.

Academic gains were also obtained:

1) The average number of correct Sullivan items per hour increased from 83 during baseline to 176 during the last fading phase; then the rate dropped to 158 during self-evaluation;

2) During the course of the 3.5 month study, a seven month gain was obtained on the Reading Vocabulary subtest of the California Achievement Test.

The Lovitt and Curtiss (1969) results received some support as the self-evaluation procedure obtained lower rates of disruptive behavior than the teacher-determined reinforcement. It remains to be seen whether the fading techniques can be effective with older and perhaps more disturbed students in psychiatric hospitals. Of course there are other populations that have not been assessed, e.g., regular class children and mildly retarded students.

Self-determined reward contingencies have been evaluated on non-academic tasks (Bandura and Perloff, 1967; Liebert et al., 1970;
Weiner and Dubanoski, 1975). Lovitt and Curtiss (1969) did use academic tasks, but they were trying to demonstrate effectiveness only; therefore, the components of the process were not evaluated. Felixbrod and O'Leary (1973) utilized arithmetic performance to investigate self-determined reward standard changes, productivity and maintenance of self- versus externally-determined reward contingencies. Many of the laboratory studies previously have made vague references to the "valuable prizes" tokens could be exchanged for and the exchange made for the first time, at the end of the study (e.g., Bandura and Perloff, 1967). To insure the effectiveness of the reinforcement procedures, Felixbrod and O'Leary displayed various prizes before and during the reinforcement conditions. The tokens were then exchanged for prizes after each trial. Twenty-four second graders were divided among three groups: a no reinforcement group; a self-determined contingency group; and a group yoked to the self-determined group. Six, 20 minute sessions were conducted on Monday, Wednesday and Friday. One at a time, the children received instructions and were given 12 sheets of arithmetic problems, 20 problems per page. Problem difficulty was equal across conditions, but difficulty increased within each session, i.e., the more problems attempted, the more difficult the problems. The self-determined group was allowed to change their schedule of reinforcement each session after the experimenter had left. The range of choices was from 10 correct problems for one point, to one correct problem for one point. All children demonstrated a significant linear trend toward more lenient reward contingencies. Self- and
experimenter-determined contingencies were equally effective in producing correct solutions and in fostering persistence on task. On both measures, the reinforcement conditions were superior to the no reinforcement group. Interestingly the percentage of correct answers was not significantly different among the three groups. Unlike Bandura and Perloff (1967) no sex effects were reported, but different sex experimenters were used: Felixbrod and O'Leary used female experimenters. Also, unsupported was the Bandura and Perloff inference that stringent self-reward criteria were consistently self-imposed. The high incentive condition in the Felixbrod and O'Leary procedure may have been a factor. Lovitt and Curtiss (1969) reported the self-selection of contingencies was superior to those teacher-selected. No support was found by Felixbrod and O'Leary, though the studies were different:

1) The Lovitt and Curtiss (1969) study took place in a classroom;

2) The Lovitt and Curtiss (1969) study utilized an applied behavior analysis design consequently the self-determined procedures were preceded by baseline and teacher-determined contingencies; the Felixbrod and O'Leary study utilized a factorial design and the subjects only experienced one experimental procedure and therefore, had no other experience with token systems.

The results of the Felixbrod and O'Leary (1973) study seem to further clarify the results of Bandura and Perloff (1967) and Glynn (1970).
The results further support the contention that self-determined contingencies are as productive as externally-determined contingencies on a meaningful task.

In marked contrast to the laboratory studies in self-control, only one study comparing the effectiveness of self- and external-control procedures in a classroom has used normal children (Glynn, 1970). The dependent measures used have not included time on task, though time in task setting was used (Felixbrod and O'Leary, 1973). Glynn, Thomas and Shee (1973), studied the effect of self-control, i.e., self-assessment, self-recording, self-determination of reinforcement and self-administration of reinforcement, on the on-task behavior of second grade New Zealand children in regular class. The children were 6-1 to 7-10 in age. The 37 children were observed daily during the 30 minute reading lesson. On-task was the percentage of 10 second observation intervals in which a child's behavior was congruent with assigned or on-going tasks. After a baseline, two different class contingencies were implemented. The first involved an experimenter sounding a clicker, after random intervals, indicating no off-task behavior had been observed. The click meant one minute of free time had been earned. A second class contingency only added additional back-up reinforcers, attractive to girls. A third identical class contingency followed a second baseline. The next (sixth) phase involved cueing. At intermittent intervals, tape recorded "beeps" were emitted. If any of the four reading groups were on-task for 10 seconds, a child recorder, publicly awarded a
point on the black board. During the seventh phase, after the "beeps" each child self-assessed and self-recorded points for being on-task at the moment of the cue. The maximum amount of reinforcement and the type of backup obtained was teacher-determined. Four weeks of no observation preceded a one week reassessment of the on-going self-control procedures. A one week baseline and one week reinstatement of the self-control procedures then completed the experiment. The results indicated that higher percentages of on-task were obtained with reinforcement and the self-control procedures resulted in the highest percentages. The class contingency percentages were 72%, 81% and 83%, respectively; the group contingency obtained 86% and the self-control procedures obtained 93%, 93%, and 90% on-task. The self-control phases had less variability compared to the others, though no test of significance was reported. Accuracy checks were made on the first self-reinforcement phase: 15% of the self-assessments and subsequent self-reward were overestimates and 9% were underestimates, the rest were accurate. The study demonstrated that regular second graders can effectively increase on-task behavior and maintain high performance with self-control procedures. Even with no penalty for inaccurate self-reward, 85% of the children in the study were accurate or underestimated the rewards earned.

Several different populations have been found to effectively utilize self-control procedures, but only after extended external-control has preceded the transition to self-control: Kaufman and O'Leary (1972) utilized 25 days of external-reinforcement; Drabman,
et al., (1973) used 35 days, and Glynn et al., (1973) used 50 days; Santogrossi et al., (1973) failed to obtain maintenance with self-reward after nine days of teacher-determined reward. Glynn and Thomas (1974) reasoned that providing cues for self-assessment and self-recording might be enough structure for normal children to use self-control procedures, and no prior experience with external-control procedures would then be required. From a class of 34 third graders, nine difficult to manage children were selected, 7-1 to 8-3 in age. The eight boys and one girl were observed for 50 minutes beginning at 9:30 every morning. On-task was the same dependent variable defined as by Glynn et al., (1973), the percentage of 10 second observation intervals in which the subject was observed on-task. Two distinct classes of behaviors were on task:

1) During teacher instruction: in seat, silent, looking at the teacher or speaker;

2) During work periods: attending to seat work.

After a 10 day baseline, during which the teacher announced the on-task requirements at the beginning of each lesson a 10 day self-control period was implemented. During the first self-control phase the subjects self-assessed and self-recorded marks if they were on-task when intermittent tape-recorded "beeps" were emitted: the time intervals varied randomly over 1, 2, 3, 4 and 5 minutes. A range of 10-12 "beeps" occurred per day. All the children in class were following these procedures to earn free time at the rate of one minute per mark. The free time was spent on the playground or in a
special room with attractive toys and activities, available at no other time. A second baseline followed and a second self-control period came next with procedures to remedy two observed difficulties:

1) The observers often noted that children were on-task for 2-3 minutes, but were off-task when a 4 or 5 minute interval elapsed; to remedy the difficulty, 4 and 5 minute intervals were excluded thus increasing the number of signals to 15-18 per session;

2) The children were also observed to have difficulty judging when they were on task, e.g., children working on seat work considered themselves on-task even when the teacher was ignored while giving additional instruction; to make the discrimination easier, a visual cue was provided: one side indicated in red lettering, look at the teacher; the other side used green lettering to describe seat work activities.

The phases effect was significant and the mean percentages of on-task were as follows: Baseline 1; 49.6; Self-control: 69.8; Baseline 2: 50.78; and Self-control and visual cue: 91.1. All subjects displayed increased variability during the first self-control phase over the first baseline. This variability during the self-control and visual cueing was less than the first baseline. The authors reported informal assessments indicating academic performance was enhanced. The study did demonstrate that with normal children extensive external-control procedures did not have to precede the
transition to self-control as others had found (e.g., Drabman et al., 1973). Glynn and Thomas have apparently identified very important variables in the acquisition of self-control behavior, i.e., cues regarding what behavior is to be reinforced and when.

Self-control has been shown to be an effective intervention technique and no less effective than externally-controlled reinforcement, but how effective is it compared to other intervention techniques? Stamps (1973) attempted to answer this question by comparing self-reward procedures to group therapy. Several personality measures made up the bulk of the dependent measures: hostile press, as measured on six Thematic Apperception Test cards; level of aspiration, i.e., number of arithmetic problems estimated to be correctly performed; confirming interval, i.e., the range from the number the subject would be surprised to get correct, to the number the subject was not surprised to get correct; need achievement; test anxiety; and arithmetic performance. The authors wanted to compare short-term group therapy, using procedures described by Rogers (1951) and Birney, Burdick, and Tevan (1969), to self-reinforcement. After pre-testing, 60 male, fourth, fifth and sixth graders were assigned to one of three groups: self-reinforcement; group therapy; and a no treatment control. The self-reinforcement procedures were administered to groups of 10 during 10, half-hour sessions. The boys were informed they could self-reward one token per page of arithmetic problems when the self-selected criterion was equalled or exceeded. The problems were self-corrected with answer sheets. The level of aspiration could only be
selected at the beginning of each session, it could not be changed during a session. The tokens were exchanged for prizes of different value, so the more tokens the more valuable the prize. The children also self-recorded self-reinforcement, so that unknown to the subject, undeserved self-reward could be evaluated. The group therapy subjects were initially given the same first arithmetic problems and asked to estimate level of aspiration and confirming intervals. No problems were attempted, though. The therapy was centered around the use of various competitive table games. The experimenter interacted with each child during each session, providing positive interaction for success and neutral interaction for failure with verbal rewards contingent on effort. Clarification and reflection of feelings associated with success or failure were also utilized. The author hypothesized that these techniques could affect a redefinition of failure for the boys. The control group was given the same first arithmetic problems with the attendant measures, but the problems were not attempted. The children were told the period was for study. All were post-tested with the TAT and Test Anxiety Questionnaire.

The self-reinforcement group and the group therapy children were both significantly different from the control on hostile press scores. The self-reinforcement group was superior to the group therapy and control conditions in reducing level of aspiration. There were no significant differences among the three groups in confirming interval, need achievement, and test anxiety. The self-reinforcement group did significantly increase arithmetic performance, but no other group
utilized daily arithmetic activities. The self-reinforcers also self-administered significantly more reinforcers that earned, so the rewards were apparently highly valued. The study did demonstrate that in two weeks both intervention techniques were able to reduce fear of failure, i.e., hostile press. The self-reinforcement group also reduced the level of aspiration to a more realistic level. The mean width of the confirming interval changed from 3.35 to 1.94 for the self-reinforcement group and it gained in size for the other groups, but the differences were not significant. Personality variables appear to be quickly influenced by self-reinforcement procedures. The lack of a control for the arithmetic activities does reduce the strength of inferences regarding self-reward.

Relatively few studies have investigated the influence of self-control procedures on personality variables as did Stamps (1973). Laboratory studies have produced some results which are relevant. The relationship between antecedent success and failure experiences may be related as children with a preponderance of either class of experiences are expected to have quite different levels of self-esteem. The previously reported laboratory studies indicated that contingent self-reward was depressed following failure and facilitated by success (e.g., Bandura and Whalen, 1964; Masters, 1972). This would logically support the expectation that children with high self-esteem would self-reward at high levels and lower self-reward would characterize children with low self-esteem. This was the case in the study by Reschely and Mittman (1973).
The relationship between verbal self-reward and self-concept and other measures was studied by Felker and Thomas (1971). White fourth graders were assigned to two groups: one group ordered a list of nine statements which were "Good to say to myself while doing schoolwork." After the evaluative statements were ordered, positiveness score was obtained providing a measure of self-initiated self-praise. The higher the score the more positive the self-praise. Prior to any of the procedures, all the subjects were pre-tested with a self-concept scale, a measure of verbal fluency, and a measure of locus of control: Intellectual Achievement Responsibility Questionnaire.

The experimental subjects were asked to read a story describing a same sex fourth grader who was described saying things to him/herself. The subject was then shown the list of nine statements and asked "Which of these things would you say to yourself, if you were Jane (John)?" after each of 10 spelling words were presented. The positiveness scores provided an after test score. The locus of control measure provided a responsibility for success score and a responsibility for failure score. The total sample obtained a significant positive correlation between self-concept and responsibility for success and this relationship held for the girls, but not the boys. A significant relationship was found between self-concept and responsibility for failure for the boys. Children with high self-concepts selected more positive statements, but no relationship was found between self-concept and the after test scores. The authors inferred that the positiveness scores, which supposedly were measures of what the subjects
would say to themselves during schoolwork, were valid indications that self-praise was related to self-concept during the task, but not after the task.

The Felker and Thomas (1971) study was replicated (Felker and Stanwyck, 1971) with a smaller sample of white fourth graders: 88 as opposed to 131. The subjects read a short story that described a same sex fourth grader. Ten spelling words were presented to the subject. After each spelling word, the nine self-evaluative statements were presented and the experimenter said "Which of these things would John (Mary) say to himself, if he were you?" The after test scores were correlated with the self-concept scale, administered before the experimental procedures. Children with high self-concepts chose more positive self-evaluative statements. Another sex effect was found: boys obtained a significant positive correlation between self-concept and self-praise; the relationship was not significant for girls. There was no significant relationship between self-concept and spelling performance. The lack of a relationship between self-concept and performance supported results regarding self-criticism by Stouwie et al., (1970). The conflicting results between the two Felker studies regarding self-concept and self-praise are difficult to evaluate. The experimental procedures were the same as were the samples. The pre-testing and sex of experimenters were apparently different and may have influenced the results. The study requires replication.

A direct modification of self-concept was attempted by Krop, Calhoon, and Verrier (1971). Emotionally disturbed children residing
in a psychiatric hospital were assigned to one of three groups of 12. The average age was 10-6. An overt reinforcement group received a token and a gum drop after making a response associated with a positive self-concept. A control group received no reinforcement. A third group was guided to use covert reinforcement: a pleasant imaginary scene was elicited after each response which was associated with a positive self-concept. The subjects were administered a self-concept scale 24 hours before the treatment, immediately after the treatment and two weeks later. A very powerful effect was measured resulting from the covert self-reinforcement. The covert condition was significantly more positive than the control group. The follow-up results were still significantly different after two weeks. Overt and covert reinforcement were not significantly different, but only the covert reinforcement was significantly different from the control group.

Though only a few published studies have focused on personality variables, a measurable effect on self-concept can be obtained in very short periods of time: one session Krop et al., 1971; and two weeks, Stamps, 1973. Self-concept generally seems to be positively related to level of self-praise, but boys and girls may be differentially affected (Felker and Stanwyck, 1971; Felker and Thomas, 1971).

A laboratory study on the characteristics of academic tasks is very relevant. One of the factors that has made the applied studies with self-control important, has been their use of meaningful tasks. Masters and Christy (1974) are apparently the first to examine the
effects of task variables on self-reinforcement. A pilot study was implemented to determine second grade children's judgments of difficulty and length of time required to complete three tasks. These tasks were then utilized in the study with a same aged sample. The tasks were card sorting of complex forms, arithmetic operations, and a form matching task, e.g., draw a line through all the forms that match the sample. All three tasks were varied according to hard-easy and long-short. There were twelve cells, four for each task; long-difficult; lone-easy; short-difficult; short-easy. The range of mean length was 95-35 seconds. The different versions were not independent as would have been ideal. Thirty-two children were administered the four versions of the three tasks. The children were told some of the tasks were easier than others and some took longer. Self-reward instructions indicated the opportunity to self-administer tokens, and the more tokens the more valuable the prize. After each task, the experimenter spread 10 tokens on the table and instructed the child to pick up as many as were deserved. The tokens were then stored for later redemption. Both length of task and difficulty of task proved to be significant on time to completion, i.e., longer tasks and difficult tasks took longer to complete. Also, children self-rewarded more after long tasks, but the difficulty dimension interacted with self-reward performance. There was no significant effect of difficulty on self-reward on the form matching; on the card sorting, difficult tasks resulted in high self-rewarding, but easy arithmetic computations resulted in high self-reward. The effect of length upon self-reward seems independent of the task, but the effect of difficulty
seems task specific. The authors suggested this phenomenon may be a result of cues regarding the quality or accuracy of the child's performance. The arithmetic and form matching tasks were visible on the table after completion. The children could then review and self-reward according to accuracy. This may have been the case with arithmetic, as the easier problems obtained higher self-reward. With these cues absent then difficulty may be the prime criterion for self-reward as the card sorting seemed to show. The authors did not explain the intermediate position of the form matching, i.e., the easy tasks obtained more self-reward, but not significantly so. This still may be congruent with the author's hypothesis, if the level of accuracy was lower than the easy arithmetic problems, then the children would have judged few to be correct. The form matching accuracy was unreported so no inference can be made.

Two applied studies used arithmetic performance as a dependent measure (Felixbrod et al., 1973; Stamps, 1973). The level of difficulty and length on task were equal between the reinforcement conditions in the Felixbrod et al. study. Stamps (1973), however, used different tasks in the experimental conditions: arithmetic with self-reward, and competitive games with group therapy. The possible difference in difficulty may have affected the level of aspiration. The children controlled the reward criterion in the self-reward group; the level of aspiration might have been lowered to be congruent with the level of accuracy. Praise in the group therapy condition was administered by the experimenter and the children in group therapy may have seen no relationship between task accuracy and reinforcement.
This may well be a factor in the self-reward group's significant decrease in level of aspiration and the maintenance of a high level of aspiration in group therapy.

The results of studies in applied settings have consistently found that self-determined reward is at least as effective as externally-determined reward (Bolstad and Johnson, 1972; Felixbrod and O'Leary, 1973; Glynn, 1970; Lovitt, 1969; Lovitt and Curtiss, 1973). If the antecedent experiences with reward systems are limited, self-determined reward has been ineffective (Santogrossi et al., 1973), but self-control procedures have proved effective with no previous training provided additional structure is added (Glynn et al., 1974). The identification of relevant cues to increase the effectiveness of self-control by Glynn et al., (1974) may be very important in the continued identification of process variables necessary to utilize self-control procedures with varied populations and across different settings.

Some studies have sampled psychiatric populations, Kaufman and O'Leary, 1972; Santogrossi et al., (1973) learning and behaviorally disordered (Drabman et al., 1973; Lovitt, 1969; Lovitt and Curtiss, 1973), disruptive children in regular classes (Bolstad and Johnson, 1972) and regular class students (Felixbrod and O'Leary, 1973; Glynn, 1970; Glynn and Thomas, 1974; Glynn et al., 1973), and have found success with self-control procedures except for one (Santogrossi et al., 1973).

The Santogrossi et al., (1973) study raised several pertinent issues. Self-evaluation was not effective in maintaining low levels of disruptive behavior with adolescent psychiatric patients as
demonstrated by Kaufman and O'Leary (1972). The Kaufman and O'Leary (1972) study was investigating reward and response cost, so the activities preceding the successful self-evaluation were different from the Santogrossi et al. (1973) self-evaluation and experimenter reinforcement activities. The experimental activities of both studies did involve reward for low rates of disruptive behavior and a basic difference was Kaufman and O'Leary's 25 days of reinforcement prior to self-evaluation compared to 15 days of prior reward for Santogrossi et al. One result has been that self-determined reward has not been demonstrated to be as effective as externally-determined reward with this population. Another issue is the term self-evaluation. Self-evaluation seems to be defined differently by different researchers. Kaufman and O'Leary (1972) and Santogrossi et al. (1973) seemed to define self-evaluation as a student's publicly stating how many points have been earned by behaving according to the rules, with no established reinforcement schedule. Kaufman and O'Leary (1972) provided tokens based on the subjects' self-evaluation. Santogrossi et al. (1973) did not. Glynn et al. (1973) has defined three different procedures that Kaufman and O'Leary (1973) and Santogrossi et al. (1973) have apparently combined in different combinations:

1) Self-assessment--the subject's examination of his own behavior to decide whether a specific behavior or class of behaviors have been performed;

2) Self-determination of reinforcement--the subject's determination of the nature and amount of reinforcement that shall be
received contingent upon the performance of a behavior or class of behaviors;

3) Self-administered reinforcement--the subject's dispensing reinforcement to himself contingent upon the performance of a class of behaviors or a specific behavior.

The children in the Kaufman and O'Leary (1972) and Santogrossi et al., (1973) studies indicated how many points they had earned during the self-evaluation phase which seems to include procedures one and two; Kaufman and O'Leary (1972) included procedure three as part of self-evaluation also. Kaufman and O'Leary (1972) have concluded that self-evaluation can maintain low levels of disruptive behavior, but Santogrossi et al., concluded that self-evaluation is ineffective; both are describing different procedures. This confusing use of a term, defined differently, seems all too prevalent.

An important finding by Drabman, et al. (1973) indicated that fading may be very important in teaching self-evaluation to behaviorally disordered children and by juxtaposing a non-reward period to reward periods obtained generalization of self-reward for the first time. A positive influence of self-reinforcement was noted by Bolstad and Johnson (1972). The non-reinforced control children in the same room with children self-rewarding had significantly lower levels of disruptive behavior than a control group not exposed to experimental subjects.

The studies on self-concept have found self-reward to be directly and positively related, i.e., children with high self-concepts
self-reward more than children with low self-concepts on imaginary school work (Felker and Thomas, 1971) and spelling tests (Felker and Stanwyck, 1971). The studies by Felker and his colleagues obtained sex differences that may well be related to sex of experimenter, i.e., no significant relationship between self-concept and self-reward was found on the spelling task for girls and apparently a male experimenter was used.

Very little attention has been given to personality variables such as self-concept. Fear of failure and level of aspiration have been studied and were found to be influenced relatively quickly (Stamps, 1973). These variables were affected as much by self-reward as by group therapy techniques, with a more realistic level of aspiration as a result of only the self-reinforcement procedures.

More attention to task characteristics should increase the meager amount of knowledge regarding self-control in the natural environment. Masters and Christy (1974) found that length of task and difficulty are two significant parameters, but independent. Difficulty seems to be task specific and strongest when cues as to accuracy and quality are absent. Much more research is needed on different tasks so that better control may be achieved to evaluate self-control.

Another example of the need for better control is the attempt at equalization of amount of reinforcement. Bolstad and Johnson (1972) did not equalize the amount of reward in both experimental conditions: the self-reinforcement group gave themselves an average of two points more per day, 7.4 compared to 5.3, but were not significantly different
from the experimenter-reinforced group. The Krop et al., (1971) study administered covert self-reinforcement and experimenter-reinforcement, but only the covert reinforcement obtained significant changes in self-concept; were the two levels or kinds of reinforcement equivalent? Masters and Mokros (1974) have discussed the scaling difficulties in the research on self-reward, but it seems to be a problem in all of the self-control research.

Summary

Unlike previous learning research, the reinforcement or punishment procedures in self-control studies are different; therefore, it is difficult to compare external-criticism to self-criticism, experimenter-reinforcement to self-reinforcement or vicarious reinforcement and overt reinforcement to covert reinforcement without some validation that the compared experimental procedures were essentially equivalent. The equivalence has not been established and the present findings in self-control research will be clearer when the scaling of reinforcement and punishment is completed.

No less important is the need for clearer specification of procedures and a more careful use of terms. For example, the results of the Krop et al., (1971) study are compelling, but the activities during which the overt and covert reinforcements occurred were unspecified. The previous discussion on the different definitions of self-evaluation (Kaufman and O'Leary, 1972; Santogrossi et al., 1973) is an example of different definitions for an often used term. A much more important variation in definitions is exemplified by
self-reinforcement. The research on social comparison processes often directs the subject to self-administer tokens from a limited number, others self-administer one token upon meeting criterion and several studies have utilized the self-administration of points (a limited number) as self-reward. Bandura (1971) cogently suggests that to the extent the subject does not have free access to the reinforcers, the procedure is a variation of external-reinforcement. If this is true, much of the present research in self-control would be considered more a kind of mutual control. Word play is not the intent here, but word play is easily a result of unclear terms.

The preponderance of published studies are laboratory research. A heavy reliance on experimental and comparative designs has done much to explicate variables and to begin the systematic study of their relationship to self-control. Perhaps this has helped to continue the use of ambiguous tasks. Work now has begun with meaningful tasks and indeed a much needed study of task characteristics has also been initiated. Much more research in applied settings is needed to see if the laboratory tested procedures will work in the natural environment. Much more use of applied behavior analysis should be helpful. As Jeffery (1974) and Mahoney (1972) have indicated, multiple baseline designs may be the most useful since no reversal of treatments is required. Therefore, behaviors such as academic achievement and aggression would not be required to return to baseline.

Several areas have not received the attention that self-reward has. Self-criticism and self-punishment have not received the
interest as in the early sixties when Aronfreed was publishing (e.g., Aronfreed, 1963). Bandura (1969) indicated that covert processes are probably more important than external processes in behavior change, yet relatively few studies have evaluated cognitive mediation. This is especially true of the research in applied settings. The evaluation of self-control in the natural environment will no doubt receive much more attention. Psychiatric populations have received little attention and no studies were found with the mildly retarded. Much more needs to be done to find effective ways to combine self-assessment, self-recording and self-administered consequences with stimulus control and cueing to design treatments which are effective with different populations at different ages. Now that self-control procedures have been demonstrated to be effective, their durability, maintenance and generalization will require more investigation.
Unlike the majority of the reviewed studies in self-control, this study was implemented in a classroom for exceptional children. Behavior measures were obtained daily and analyzed through the use of a reversal design. The results were evaluated for the individual and group behavior graphically and statistically. A detailed description follows.

Subjects

The experimental subjects were selected from a learning disabilities/behavior disorders (LD/BD) class. The class was selected in cooperation with the LD/BD Program supervisor, the building principal and the classroom teacher. The criteria used in the class selection were as follows:

1. The principal and teacher must volunteer to participate;
2. The teacher must have at least one year of teaching experience;
3. The teacher must be judged competent by the principal and the LD/BD supervisor;
4. The class census must be stable, i.e., no transfers of more than three children will be anticipated;
5. On-task behavior must comprise no more than 75% of the experimental period.

The class included two girls and eight boys between six years, eight months and seven years, ten months of age. All the children had received psychological and medical evaluations attesting to the appropriateness of the placement. The children all obtained I.Q.'s above 80 on individually administered tests.

During a seven day baseline period, seven of the ten children in the class were found to be on-task less than 75% of the experimental period. Those seven children were selected as the experimental population.

**Setting**

The study took place in a suburban-rural public school system of approximately 16,000 children. The LD/BD class selected was in a building of approximately 750 children with grades kindergarten through fifth. Two LD/BD classes were housed in the building.

Academic instruction in the classroom was highly individualized with a combination of teacher-made materials and manufactured or printed instructional materials. The experiment took place during a forty-five minute period of reading instruction which was fifty to sixty minutes in length. Typically, the teacher instructed the children to practice memorizing words listed on the blackboard. After studying the words, the children orally read the words to the teacher and then silently read an assigned story, twice. Then work book or other pencil paper activities related to the reading instruction
followed. While the preceding activities were in progress, the teacher called small groups of the children to a table at the side of the room to read aloud the story they had been assigned to read silently. The experimental period began when the teacher initiated instructions for the reading activities.

**Materials**

The visual cue was a one-quarter inch cardboard sign of 432 square inches (18" x 24"). One side said "Look and Listen" in four-inch orange letters on a white background. The other side was divided in half by a yellow diagonal. The left upper half said "seatwork," the right lower half said "raise hand," in blue letters on a white background. Each direction was accompanied by a four-inch comic strip character exemplifying the instruction. For example, by the word "seatwork" was a picture of a boy sitting at a desk writing; by the words "raise hand" was a picture of a girl with her hand raised. The comic strip characters were copied with the aid of an opaque projector and colored with magic markers.

The auditory cue was produced by a Super SN-10 SP, Guitar pitch pipe. The cue was the reed used to tune the high E string, i.e., the highest pitched guitar string. The auditory cues were recorded on a Sony-Matic TC-66 Cassette-Corder. The one-second tone was recorded at intervals of .5, 1, 2, or 3 minutes. The intervals were sequenced using a randomized procedure (Kerlinger, 1973).

The children's self-awarded points or x's were self-recorded on blank IBM cards marked with thirty numbered spaces (Appendix A).
A card with 24 numbered spaces was used for one day only (Appendix B). Since the children seemed to be self-rewarding more frequently than they were being cued, the number of spaces were increased to 30 to evaluate this possibility. The teacher had previously utilized reinforcement systems with the class. Reinforcers and reinforcing activities were selected in cooperation with the teacher. Reinforcers used were bubble making materials, extra recess, coloring activities, crayons, money to buy pencils, erasers, access to magic markers, puzzles, coloring books, and sewing activities.

Behavior Definition and Recording

The dependent variable, on-task behavior, was the percentage of ten-second observation intervals in which an individual child's behavior was classified as on-task. On-task behaviors were selected to conform as closely as possible to those utilized by Glynn and Thomas (1974) yet still conform to the specific classroom situation encountered. Two general classifications of on-task were utilized:

1. Looking at and listening to the teacher during instructions or after the teacher said the equivalent of "Look at me";

2. Working on assignments, whether this was alone at a desk, with a partner, or at an interest center.

The specific behavior descriptions follow:

On-task: Writing, reading, looking at or manipulating materials relevant to current assignments or teacher directions. Also, talking to the teacher with permission. On-task behavior need occur for only one second to be recorded. Self-recording was
considered on-task if it occurred within 30 seconds after an auditory cue.

Off-task: Looking at something other than assignment relevant material, talking without permission or grooming for an entire ten-second interval. Also, any combination of an off-task behavior and neutral behavior.

Neutral Behavior: Hand raised, moving to or from different parts of the room, and looking through the desk. Neutral behavior occurred for only a fraction of an interval because if neutral behavior occupied the entire interval, it was considered on-task.

Accurate Self-reward: Subsequent to an auditory cue occurring between the second and ninth seconds of an interval, accurate self-reward was the recording of an x on the desk card if the child was on-task when the auditory cue occurred. The self-recording must occur within thirty seconds of the auditory cue.

Inaccurate Self-reward: Self-recording an x when off-task at the moment the auditory cue occurred or self-recording an x when no beep has occurred within the preceding thirty seconds.

Teacher Praise: The words "Right," "Correct," "Very Good," "Fine," and "Thank you" specifically, and any other words or statements when used in an approving manner regarding student behavior or school work. Due to an ambiguous use of the following terms, a decision was made to not consider them praise in this study: "Alright," "Yes" and "Uh Huh" (indicating yes). Teacher praise
was recorded continuously, but only once during each fifteen-second interval that it occurred.

It should be noted that on-task behavior had to only occur for one second to be recorded. Off-task behavior had to make up the entire interval or juxtapose neutral behavior to comprise the entire interval, in order to be recorded.

A sequential time-sampling observation procedure was utilized due to the low percentage of error obtained by this procedure compared to other time-sampling methods (Thompson, Holmberg, and Baer, 1974). The experimenter sat in the left front corner of the room relative to the direction the seated class faced. Parallel to the front of the room, the children were seated in three rows: five in the first row, three in the second and two in the third. The child seated nearest the experimenter was observed first, then the child next farthest away in that row was observed until all in row one were observed. The child nearest the experimenter in row two was observed next with the child next farthest away from the experimenter observed next. Using this near to far sequence in each row consecutively, resulted in each child being observed once every 150 seconds. The recording form (see Appendix B) allowed each child's behavior to be recorded in a column beneath their name. The recording form allowed the simultaneous recording of five behaviors:

1. On-task;  
2. Off-task;  
3. Accurate self-reinforcement;  
4. Inaccurate self-reinforcement;  
5. Teacher praise.
Each child was observed for ten seconds. At the end of the ten-second interval the observer took five seconds to circle the symbols representing any of the five behaviors that occurred. Teacher praise was recorded continuously, but only once during any fifteen second interval.

The experimenter was the observer and a volunteer observer provided reliability measures. A thirty-minute video tape of the experimental class was made prior to the beginning of the study. The experimenter and the assistant viewed the video tape to practice using the observation and recording procedures. The children viewed on the tape were observed in an agreed upon sequence. Both observers viewed the tape, stopping every five minutes to compute percentages of agreement 

\[
\frac{\text{Intervals in Agreement}}{\text{Total Intervals}} \times 100
\]

and to discuss and rectify discrepancies. Both observers were out of each other's sight during the practice observation to remove non-verbal cues. The observers continued to record the video taped behaviors until 90% agreement on on-task behavior was achieved on three consecutive five-minute viewing periods. Subsequently, the observers viewed the entire thirty-minute tape continuously and ceased practicing, since the percentage of agreement was above 85%.

**Design of the Study**

A reversal design incorporating three different experimental procedures was utilized in the study. Each of three independent variables, visual cueing alone, auditory cueing and self-reward, visual and auditory cueing with self-reward, was implemented. The most effective of the three procedures, visual and auditory cueing with self-reward,
was replicated within the reversal design to assess the effect on the dependent variable, on-task behavior. The reversal design permitted a valid evaluation of the functional relationships between the variables studied.

The reversal design is often described as an applied behavioral analysis. An increasing interest in applied behavior analysis has been manifested in recent years (Journal of Applied Behavior Analysis, 1968-1976). The recent interest in analyzing the functional relationships between variables of a single subject alone or in small groups in applied settings has largely resulted from the failure of between group designs to provide information regarding the learning process of the individual (Cooper, 1974). Kazdin (1973) has discussed the problem "that averages from group data usually have no analogue in representing the behavioral process of individuals" (p. 518). Also, it has been stated that when a statistical evaluation is based on inter-subject variability, lawful effects of variables may be obscured (Sidman, 1960).

There are several advantages in using an applied behavioral analysis. These advantages include a small number of subjects may be used, no indirect measures are required and the data are real time measures as opposed to brief post-treatment samples of behavior. Finally, applied behavior designs make replication an essential part of the design. Replication is considered the strongest evidence of a treatment effect (Edwards, 1965).

Replication may be effective in assessing the reliability and generalizability of data (Sidman, 1960). The reliability of a
treatment effect can be assessed with a single organism design. Generalizability requires intersubject replication. No inferences regarding generalizability are justified by this study and none are attempted. However, the replication of this study with different subjects does satisfy the intersubject requirement of generalization and therefore this study does add to the accumulation of scientific knowledge.

The preceding discussion was the rationale for selecting a reversal design to assess the cueing procedures in various combinations with self-reward as described in Figure 1.

**Design Conditions**

There were six different design phases in the study. Data were gathered daily throughout the study on teacher praise and on-task behavior.

**Baseline 1**

During the first baseline the teacher conducted the class as usual; the teacher gave instructions and the reading period followed. The experimenter urged the teacher to use all control methods as usual, e.g., praise or rebuke. Three behaviors were recorded daily, on-task, off-task and teacher praise. Once every three days, and at least twice during baseline 1, a second observer was present to obtain a reliability measure. After a minimum of five days and when a stable or decreasing trend in on-task behavior obtained, the first experimental procedure was implemented.
Percentage of intervals on-task

Baseline 1
Visual cueing

Baseline 2
Auditory cueing and self-reinforcement

Session: Daily experimental periods

Baseline 2
The most effective of 2, 3, or 4
**Intervention 1: Visual Cueing**

On the first day of intervention 1, the teacher showed the students the sign and then hung it at the front of the room, five feet high. The sign was suspended by yarn so it could be turned easily. The teacher told the children they were doing what one side of the sign said, i.e., to look at and listen to the teacher. The teacher then turned the sign and asked that somebody show her what the sign said. The children easily read the sign and displayed no difficulty in showing her what the sign described. During the next two days, at the beginning of the period, the teacher reminded the children that the sign was to help them. During the period, the teacher often asked the children if they were doing what the sign said. The teacher manipulated the sign so that when she wanted all the children's attention, the proper side of the sign was showing. Otherwise, the "seatwork" "raise hand" side was displayed. At least once during each period the teacher required everybody's attention.

On-task, off-task and teacher praise were recorded daily. Once every three days, and at least twice during this phase, a second observer was present to obtain a reliability measure. After a minimum of five days, and after a stable trend of on-task behavior was obtained, intervention 2 was implemented.

**Intervention 2: Auditory Cueing and Self-reinforcement**

Prior to the experimental period, the teacher drew on the blackboard a figure similar to the Self-Recording Strip (see Appendix A)
that had already been taped to the children's desks. The teacher explained that if the children heard a beep (auditory cue) while they were doing what they were supposed to do, they could make an x on their card on their desk. She demonstrated making the x, then turned on the tape recorder and after the beep occurred, she instructed the children to make an x on the card. The teacher explained that the x's could be exchanged between 2:00 and 2:30 for a surprise. If a child got more than eighteen x's then they would get a surprise. After the first day the criterion for getting the surprise was 20 x's and whoever had the largest number of x's got the surprise first, with the next highest number receiving the surprise next. All earned points were expended daily.

The children were told to only mark their card if they were doing what they were supposed to only when they heard a beep. They were urged to be honest and not to talk about anybody else's card. Occasionally, the teacher saw a child about to mark their card when they had been off-task; she told them not to mark their card and why. The teacher occasionally praised accurate self-awards.

The auditory cues occurred at random intervals of .5, 1, 2 or 3 minutes. On-task, off-task, accurate self-reward, inaccurate self-reward and teacher praise were recorded. A second observer was present once every three days and at least twice during intervention 2. After a stable or declining trend of on-task behavior was obtained, intervention 3 was implemented.
Intervention 3: Visual and Auditory Cueing With Self-Reward

On the first day of intervention 3, the teacher repeated the previous instructions on how to use the sign and she reminded the students about being honest when giving themselves x's. She emphasized that the sign, the beeps and the fun things were to help the children learn. The teacher then manipulated the sign to correlate with the classroom activities, i.e., when she wanted the attention of all the children, the side of the sign saying "Look and Listen" was displayed, otherwise the other side was displayed, "raise hand" and "seat work."

On-task, off-task, accurate self-reward, inaccurate self-reward and teacher praise were recorded. Once every three days, and at least twice during intervention 3, a second observer was present to obtain reliability measures. After a minimum of five days and after a stable or increasing rate of on-task behavior was obtained, baseline 2 was implemented.

Baseline 2

All cueing and self-reward were removed during this phase. The teacher informed the children that it was no longer necessary for the children to have the aids and extra things which had been present before. The class continued as before, but without the visual or auditory cueing, and with no self-rewarding and no backups during the 2:00-2:30 period.

On-task, off-task and teacher praise were recorded during this phase. Once every three days, and at least twice during baseline 2, a second observer was present to obtain reliability measures. After
a minimum of five days and after a stable or declining percentage of on-task behavior was obtained, intervention 4 was implemented.

**Intervention 4: Visual and Auditory Cueing With Self-Reward**

The final intervention selected obtained the highest percentage of on-task behavior among the three interventions. The instructions on the use of the sign, the beeps and the self-awarding of x's were presented by the teacher as before. The teacher informed the children that fun things would be available between 2:00 and 2:30 and the child with the most x's would get first choice, the child with the second most second choice and so on. The teacher manipulated the sign to correlate with classroom activities.

Once every three days, and at least twice during intervention 4, a second observer was present to obtain reliability measures. After a minimum of five days and after a stable trend in on-task behavior was obtained, intervention 4 was terminated. On-task, off-task, accurate self-reward, inaccurate self-reward and teacher praise were recorded.

**Reinforcement Period**

At 2:00, the teacher asked the children to put away their work. Then the teacher called the children back to her desk to select activities for the next thirty minutes or occasionally to select an object to keep, e.g., a tablet or eraser. The child with the most x's was called first, the child with the next most x's was second and so on. When there were ties, all the children with the same number of x's
were considered first, but the teacher chose the student who had not
been first for the longest period of time to select first. Every
child was offered something or some activity each day. The most
attractive backups were offered first to increase the incentive value
of the x's.

**Data Collection and Analysis**

Data for on-task, off-task, accurate self-reward, inaccurate
self-reward and teacher praise were recorded on recording forms
(see Appendix C). For each child, the number of on-task intervals
recorded were divided by the total number of intervals to obtain a
daily percentage of on-task behavior. The number of correct self­
reinforcements recorded were divided by the total number of self­
reinforcements to obtain a daily percentage of accurate self-reinforce­
ments when self-reinforcement was part of an intervention. The daily
percentages of on-task and accurate self-reward were graphed for each
child. The daily frequency of teacher praise was graphed for each
day of the study. The total daily percentages of on-task and accurate
self-reward were graphed for the entire experimental group also.

Total percentages of on-task behavior for the entire experimental
population were analyzed using a distribution-free test for ordered
alternatives based on Friedman rank sums (Hollander and Wolfe, 1973).

A second observer was trained to make the observations necessary
for the study until an average percentage of agreement with the exper­
imenter above 85% was obtained for on-task behavior. A reliability
measurement between the two observers was made at least once every
three days during each phase of the study. When the interobserver reliability was below 80%, the training procedures were re-implemented until the practice agreement percentages averaged at least 85%. On a few occasions the behavior definitions were altered slightly to remove ambiguous or difficult to label behavior.

Interobserver reliability measures were computed for each behavior recorded. The number of agreements between observers were divided by the total number of observations and multiplied by 100 to obtain a percentage of agreement (Wasik, Senn, Welch and Cooper, 1969):

$$\frac{\text{Number of Intervals in Agreement}}{\text{Total Number of Intervals}} \times 100 = \% \text{ of Agreement}$$

In order to reduce the possibility of high percentages of agreement which were an artifact of high or low rates of behavior (Bijou, Peterson, Harris, Allen, and Johnston, 1969), additional agreement percentages were computed. The percentage of agreement for scored intervals only, allowed an accurate assessment of observer reliability at low frequency levels (Hawkins and Dotson, 1975). This statistic was computed by dividing the number of scored intervals in agreement for a behavior by the total number of scored intervals for that behavior, multiplied by 100. The scored interval percentage of agreement was computed for each behavior recorded during each reliability check. Also, the unscored intervals percentage of agreement were computed to provide a more accurate assessment of observer reliability for frequently occurring behavior also (Hawkins and Dotson, 1975). The unscored interval agreement percentage computed by dividing the number of unscored intervals in agreement by the total number of unscored
intervals for that behavior, multiplied by 100. If the total percentage of agreement of scored and unscored intervals for a behavior fell below 80%, the training procedure was reinstated until the average was 85% for the 30 minute training video tape.

In order to provide another estimate of the validity of the observations the behavior measurements for each behavior by the second observer were graphed with those data obtained by the experimenter. The graphing of both observers' recordings allows inferences regarding any systematic biases on the part of either observer and the direction of any observed biases.

Interobserver reliability measures were computed for total observations for each behavior during each phase of the study and for the total study. Scored interval and unscored interval percentages of agreement were computed for each behavior during each phase of the study and for the total study. The measures obtained by the second observer were graphed with the experimenter's observations to allow inspection for biases and directionality among any inferred biases.
CHAPTER IV
PRESENTATION AND ANALYSIS OF DATA

The effects of the treatment are presented for analysis individually for each child and for the entire group. A graphic presentation of individual and group data are followed by a statistical analysis of the treatment effects upon the children as a group. A discussion of the data pertaining to each question answered by the study is presented last.

Inter-Observer Agreement

Inter-observer agreement of data recording was determined during each condition of the study. A reliability measurement with a second observer was obtained once every three days and at least twice during each condition of the study. Two reliability measures were obtained to more accurately assess the reliability of low frequency and high frequency behaviors. Low frequency behaviors are more accurately assessed by scored interval reliability and high frequency behaviors are more accurately assessed by unscored interval reliability. The statistic is computed by dividing the agreements by the summed agreements and disagreements, multiplied by 100.

The inter-observer agreement percentages for on-task, teacher praise, accurate self-reinforcement and inaccurate self-reinforcement are presented in Tables 1 through 4, respectively. Table 1 presents
### Table 1
Inter-Observer Agreement Percentages For Intervals

On-Task By Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Measurements</th>
<th>Scored Interval Agreement</th>
<th>Unscored Interval Agreement</th>
<th>Total Percentage Per Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>2</td>
<td>89%</td>
<td>76%</td>
<td>85%</td>
</tr>
<tr>
<td>Intervention 1</td>
<td>3</td>
<td>89%</td>
<td>72%</td>
<td>85%</td>
</tr>
<tr>
<td>Intervention 2</td>
<td>5</td>
<td>91%</td>
<td>68%</td>
<td>86%</td>
</tr>
<tr>
<td>Intervention 3</td>
<td>3</td>
<td>88%</td>
<td>62%</td>
<td>82%</td>
</tr>
<tr>
<td>Baseline 2</td>
<td>2</td>
<td>90%</td>
<td>66%</td>
<td>85%</td>
</tr>
<tr>
<td>Intervention 4</td>
<td>2</td>
<td>89%</td>
<td>63%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Total Percentage of Agreement 90% 72% 84%

### Table 2
Inter-Observer Agreement Percentages

For Teacher Praise By Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Measurements</th>
<th>Scored Interval Agreement</th>
<th>Unscored Interval Agreement</th>
<th>Total Percentage Per Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>2</td>
<td>13%</td>
<td>91%</td>
<td>84%</td>
</tr>
<tr>
<td>Intervention 1</td>
<td>3</td>
<td>45%</td>
<td>93%</td>
<td>88%</td>
</tr>
<tr>
<td>Intervention 2</td>
<td>5</td>
<td>55%</td>
<td>95%</td>
<td>91%</td>
</tr>
<tr>
<td>Intervention 3</td>
<td>3</td>
<td>59%</td>
<td>94%</td>
<td>90%</td>
</tr>
<tr>
<td>Baseline 2</td>
<td>2</td>
<td>68%</td>
<td>97%</td>
<td>95%</td>
</tr>
<tr>
<td>Intervention 4</td>
<td>2</td>
<td>23%</td>
<td>92%</td>
<td>86%</td>
</tr>
</tbody>
</table>

Total Percentage of Agreement 48% 94% 89%
Table 3
Inter-Observer Agreement Percentage for Accurate Self-Reinforcement By Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Measurements</th>
<th>Scored Interval Agreement</th>
<th>Unscored Interval Agreement</th>
<th>Total Percentage Per Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention 2</td>
<td>5</td>
<td>92%</td>
<td>98%</td>
<td>96%</td>
</tr>
<tr>
<td>Intervention 3</td>
<td>3</td>
<td>52%</td>
<td>98%</td>
<td>96%</td>
</tr>
<tr>
<td>Intervention 4</td>
<td>2</td>
<td>92%</td>
<td>99%</td>
<td>99%</td>
</tr>
</tbody>
</table>

Total Percentage of Agreement 84% 98% 96%

Table 4
Inter-Observer Agreement Percentage For Accurate Self-Punishment By Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Measurement</th>
<th>Scored Interval Agreement</th>
<th>Unscored Interval Agreement</th>
<th>Total Percentage Per Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention 2</td>
<td>5</td>
<td>33%</td>
<td>98%</td>
<td>99%</td>
</tr>
<tr>
<td>Intervention 3</td>
<td>3</td>
<td>80%</td>
<td>100%</td>
<td>99%</td>
</tr>
<tr>
<td>Intervention 4</td>
<td>2</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Total Percentage of Agreement 67% 99% 100%
Table 2 reveals the inter-observer agreement percentages for teacher praise, by condition, for the children as a group. Also included are the number of measures, scored interval agreement, unscored interval agreement and total agreement percentage, by condition, for the study.

Table 3 presents inter-observer agreement percentages for accurate self-reinforcement, by condition, for the children as a group. Also presented are the number of measures, scored interval agreement, unscored interval agreement and percentage of agreement, by condition, for the study.

Table 4 reveals the inter-observer agreement percentages for accurate self-punishment, by condition, for the children as a group. Also presented, are the number of measures, scored interval agreement, unscored interval agreement and total percentage of agreement, by condition, for the study.

Presentation and Analysis of Data

To determine if a functional relationship exists between the dependent variable, on-task behavior, and the independent variables, visual cueing, auditory cueing and self-reward, and visual and auditory cueing with self-reward, the daily percentages of on-task
intervals were recorded under six conditions and analyzed.

Subject 1

Baseline: During baseline, Subject One was on-task during 72 percent of the intervals observed.

Visual Cueing: Subject One was on-task 68 percent of the intervals observed. This was a 4 percent decrease from baseline.

Auditory Cueing and Self-Reward: An increase of 10 percent, from 68 percent to 78 percent was obtained for the observed intervals. This was a 6 percent increase over baseline.

Visual and Auditory Cueing with Self-Reward: No change was obtained as Subject One continued at 78 percent on-task during this condition.

Baseline 2: Though all cueing and self-reward were removed, an increase of 14 percent to 92 percent was obtained during this condition for Subject One. This was 20 percent higher than baseline.

Visual and Auditory Cueing with Self-Reward: A decrease of 13 percent to 79 percent was obtained during the final condition for Subject One.

Plate 1 represents the daily and total percentages of intervals on-task, per condition for Subject One. Figure a represents the daily percentages per condition and Figure b represents the total percentage of intervals on-task per condition.
Plate 1. Daily and total percentage of intervals on-task per condition for Subject One.

Figure a. Daily percentage of intervals on-task per condition.

Figure b. Total percentage of intervals on-task per condition.

B1: Baseline 1  T3: Visual and Auditory Cuing
T1: Visual Cuing  B2: Baseline 2
T2: Auditory Cuing and Self-Reward  T4: Visual and Auditory Cuing

with Self-Reward
Subject 2

Baseline: During baseline, Subject Two was on-task during 57 percent of the observation intervals.

Visual Cueing: There was a 1 percent increase during the second condition as Subject Two was on-task 58 percent of the observation intervals.

Auditory Cueing and Self-Reward: An increase of 6 percent over baseline and 5 percent over Visual Cueing was obtained as Subject Two was on-task during 63 percent of the observation intervals.

Visual and Auditory Cueing with Self-Reward: An 8 percent increase to 71 percent on-task was obtained. This was a 14 percent increase over baseline for Subject Two.

Baseline 2: A 1 percent increase to 72 percent was obtained for Baseline 2 for Subject Two. This was 15 percent higher than baseline.

Visual and Auditory Cueing with Self-Reward: A decrease of 10 percent from baseline 2 down to 62 percent was obtained for Subject Two during the final condition of the study.

Plate 2 represents the daily and total percentages of intervals on-task, per condition for Subject Two. Figure a represents the daily percentages per condition and Figure b represents the total percentage of intervals on-task per condition.

Subject 3

Baseline: During baseline, Subject Three was on-task 66 percent of the observed intervals.
Plate 2. Daily and Total Percentage of Intervals On-Task by Condition for Subject Two

Figure a. The daily percentage of intervals on-task.

Figure b. The total percentage of intervals on-task for each experimental phase.

- $B_1$: Baseline
- $I_1$: Visual Cuing
- $I_2$: Auditory Cuing and Self-Reward
- $I_3$: Visual and Auditory Cuing with Self-Reward
- $I_4$: Visual and Auditory Cuing with Self-Reward
- $B_2$: Baseline 2
Visual Cueing: A 22 percent increase to 88 percent on-task was obtained during this condition for the observed intervals.

Auditory Cueing and Self-Reward: A 2 percent increase to 90 percent on-task was obtained for Subject Three. This was a 24 percent increase over baseline.

Visual and Auditory Cueing with Self-Reward: A drop of 6 percent to 84 percent on-task was obtained for Subject Three during this condition.

Baseline 2: A drop of 3 percent to 81 percent on-task was obtained during this condition. This was 15 percent higher than baseline.

Visual and Auditory Cueing with Self-Reward: A 4 percent increase to 85 percent intervals on-task was obtained for Subject Three.

Plate 3 represents the daily and total percentages of intervals on-task, per condition for Subject Three. Figure a represents the daily percentages per condition and Figure b represents the total percentage of intervals on-task per condition.

Subject 4

Baseline: Subject Four was on-task 65 percent of the observation intervals.

Visual Cueing: A 15 percent increase to 80 percent on-task was obtained for Subject Four.

Auditory Cueing and Self-Reward: An 8 percent increase to 88 percent on-task was obtained. This was 23 percent higher than baseline.
Plate 3. Daily and Total Percentage of Intervals On-Task by Condition for Subject Three

Figure a. The daily percentage of intervals on-task.

Figure b. The total percentage of intervals on-task for each experimental phase.

B₁: Baseline
I₁: Visual Cuing
I₂: Auditory Cuing and Self-Reward
I₃: Visual and Auditory Cuing
I₄: Visual and Auditory Cuing with Self-Reward

* = Field trip
Visual and Auditory Cueing with Self-Reward: Subject Four was on-task 71 percent of the observed intervals. This was a decrease of 18 percent from the previous condition.

Baseline 2: A 2 percent increase to 73 percent on-task was obtained for the observed intervals for Subject Four.

Visual and Auditory Cueing with Self-Reward: a 15 percent increase to 87 percent on-task was obtained for Subject Four. This was a 22 percent increase in on-task intervals over baseline.

Plate 4 represents the daily and total percentages of intervals on-task, per condition for Subject Four. Figure a represents the daily percentages per condition and Figure b represents the total percentage of intervals on-task per condition.

Subject 5

Baseline: Subject Five was on-task 58 percent of the observed intervals during baseline.

Visual Cueing: A 14 percent increase to 72 percent on-task intervals was obtained by Subject Five.

Auditory Cueing and Self-Reward: A decrease of 3 percent to 69 percent on-task was obtained for Subject Five. This was 11 percent higher than baseline.

Visual and Auditory Cueing with Self-Reward: An 8 percent increase to 76 percent on-task intervals was obtained by Subject Five for this condition. This was 18 percent higher than baseline.
Plate 4. Daily and Total Percentage of Intervals On-Task by Condition for Subject Four.

Figure a. The daily percentage of intervals on-task.

Figure b. The total percentage of intervals on-task for each experimental phase.

- $B_1$: Baseline
- $I_1$: Visual Cueing
- $I_2$: Auditory Cueing and Self-Reward
- $I_3$: Visual and Auditory Cueing
- $I_4$: Visual and Auditory Cueing with Self-Reward

* = Field trip
Baseline 2: A 1 percent decrease to 75 percent on-task intervals was obtained by Subject Five. This was 17 percent higher than baseline.

Visual and Auditory Cueing with Self-Reward: A 2 percent increase to 77 percent on-task intervals was obtained by Subject Five. This was a 19 percent increase over baseline.

Plate 5 represents the daily and total percentages of intervals on-task, per condition for Subject Five. Figure a represents the daily percentages, per condition and Figure b represents the total percentage on-task, per condition.

**Subject 6**

Baseline: Subject Six was on-task 70 percent of the intervals observed.

Visual Cueing: An 8 percent decrease to 62 percent intervals on-task was obtained by Subject Six.

Auditory Cueing and Self-Reward: A 3 percent decrease to 59 percent on-task was obtained. This was 11 percent below baseline for Subject Six.

Visual and Auditory Cueing with Self-Reward: A 13 percent increase to 72 percent intervals on-task was obtained by Subject Six. This was 2 percent higher than baseline.

Baseline 2: A 4 percent increase to 76 percent on-task was obtained by Subject Six. This was 6 percent higher than baseline.
Plate 5. Daily and Total Percentage of Intervals On-Task by Condition for Subject Five.

Figure a. The daily percentage of intervals on-task.

Figure b. The total percentage of intervals on-task for each experimental phase.

$B_1$: Baseline
$I_1$: Visual Cueing
$I_2$: Auditory Cueing and Self-Reward
$I_3$: Visual and Auditory Cueing
$I_4$: Visual and Auditory Cueing with Self-Reward
Visual and Auditory Cueing with Self-Reward: A 1 percent increase to 77 percent intervals on-task was obtained by Subject Six. This was 7 percent higher than baseline.

Plate 6 represents the daily and total percentages of intervals on-task, per condition for Subject Six. Figure a represents the daily percentages, per condition and Figure b represents the total on-task percentage, per condition.

Subject 7

Baseline: Subject Seven obtained 69 percent intervals on-task.

Visual Cueing: A decrease of 11 percent to 59 percent on-task was obtained by Subject Seven.

Auditory Cueing and Self-Reward: An increase of 18 percent to 76 percent intervals on-task was obtained by Subject Seven. This was a 7 percent increase over baseline.

Visual and Auditory Cueing with Self-Reward: A 1 percent decrease to 75 percent intervals on-task was obtained by Subject Seven.

Baseline 2: A 10 percent increase to 85 percent intervals on-task was 16 percent higher than baseline for Subject Seven.

Visual and Auditory Cueing with Self-Reward: A 1 percent increase to 86 percent on-task was obtained by Subject Seven. This was 17 percent higher than baseline.

Plate 7 represents the daily and total on-task percentages, per condition for Subject 7. Figure a represents the daily percentages and Figure b represents the total percentage on-task, per condition.
Plate 6. Daily and Total Percentage of Intervals On-Task by Condition for Subject Six.

**Figure a.** The daily percentage of intervals on-task.

**Figure b.** The total percentage of intervals on-task for each experimental phase.

- $B_1$: Baseline
- $I_1$: Visual Cueing
- $I_2$: Auditory Cueing and Self-Reward
- $I_3$: Visual and Auditory Cueing
- $I_4$: Visual and Auditory Cueing with Self-Reward

$\# =$ Field trip
Plate 7. Daily and Total Percentage of Intervals On-Task by Condition for Subject Seven.

Figure a. The daily percentage of intervals on-task.

Figure b. The total percentage of intervals on-task for each experimental phase.

- $B_1$: Baseline
- $I_1$: Visual Cueing
- $I_2$: Auditory Cueing and Self-Reward
- $I_3$: Visual and Auditory Cueing
- $I_4$: Visual and Auditory Cueing with Self-Reward

*= Field trip
The individual total accuracy percentages for self-reward and self-punishment ranged from 55% to 77% and the total accuracy percentages for self-punishment ranged from 28% to 100%. Due to the small number occasions that self-reward and self-punishment were observed individually, these results must be interpreted with caution.

Subject One was observed accurate in self-rewarding from 60% to 85% of the events measured. Accurate self-punishment was observed to be 33%. Subject Two was accurate in self-rewarding from 25% to 100% and accurate in self-punishment from 0% to 40%. Subject Three was accurate in self-rewarding from 50% to 89% and accurate in self-punishment from 0% to 50%. Subject Four was accurate in self-rewarding from 57% to 100% and no accuracy percentage was available because no self-punishment was observed. Subject Five was accurate in self-rewarding from 62% to 73% and accurate in self-punishment 100%. Subject Six was not observed to self-punish, but was accurate in self-rewarding from 64% to 75% of the events observed. Subject Seven was accurate in self-rewarding from 73% to 100% and accurate in self-punishment from 0% to 100%. The raw data and accuracy percentages for individuals for self-reward and self-punishment are presented in Table 5. The accuracy percentages are presented by condition with total accuracy percentages for each child.

**Group Data**

The effects of the independent variables, visual cueing, auditory cueing and self-reward, and visual and auditory cueing with self-reward,
Table 5

Raw Data and Agreement Percentages Between Individuals and the Observer for Self-Reward and Self-Punishment, By Condition

<table>
<thead>
<tr>
<th>Children</th>
<th>$I_2$</th>
<th>$I_3$</th>
<th>$I_4$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+S-R</td>
<td>-SR</td>
<td>+SR</td>
<td>-SR</td>
</tr>
<tr>
<td></td>
<td>Data %</td>
<td>Data %</td>
<td>Data %</td>
<td>Data %</td>
</tr>
<tr>
<td>1</td>
<td>11/13 = 85%</td>
<td>1/3 = 33%</td>
<td>3/5 = 60%</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2/8 = 25%</td>
<td>0/2 = 0%</td>
<td>6/6 = 100%</td>
<td>2/5 = 40%</td>
</tr>
<tr>
<td>3</td>
<td>8/9 = 89%</td>
<td>1/3 = 33%</td>
<td>5/8 = 62%</td>
<td>3/6 = 50%</td>
</tr>
<tr>
<td>4</td>
<td>2/2 = 100%</td>
<td>0/2 = 0%</td>
<td>4/7 = 57%</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>7/10 = 70%</td>
<td>2/2 = 100%</td>
<td>11/15 = 73%</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>7/11 = 64%</td>
<td>0/2 = 0%</td>
<td>3/4 = 75%</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>11/15 = 73%</td>
<td>1/1 = 100%</td>
<td>8/11 = 73%</td>
<td>0/1 = 0%</td>
</tr>
</tbody>
</table>

The ratios of correct observations divided by the sum of correct plus incorrect and the equivalent percentage for accurate self-reward and self-punishment for each child, by condition.

$I_2$ = Auditory cueing with Self-Reward. $I_4$ = The second time visual and auditory cueing were introduced.

+SR = Self-Reward. -SR = Self-Punishment. 0 = No events were observed.
upon on-task behavior were analyzed by examination of the daily and total percentages of intervals on-task, per condition, for all seven children. The total percentage of intervals on-task are presented in Figure 2. Figure a represents the daily percentages of intervals on-task, per condition, for all seven children. Figure b represents the total percentage of intervals on-task, per condition for all seven children.

The effects of teacher praise upon the daily percentages of intervals on-task were analyzed by examination of the daily frequencies of teacher praise per condition and the examination of the daily percentages of intervals on-task, per condition for all seven children. The daily frequencies of teacher praise and the daily percentages of intervals on-task are presented in Figure 3. Figure a represents the daily frequencies of teacher praise and Figure b represents the daily percentages of intervals on-task.

The accuracy of self-evaluations were considered the same as the accuracy of self-reward and was analyzed by computing accuracy percentages. The number of accurate self-rewards was divided by accurate plus inaccurate self-rewards and multiplied by 100. The raw data and percentages of accurate self-reward by condition, for the group of seven children are as follows: auditory cueing with self-reward (Intervention 3), 36/50 = 70%; and total accuracy of self-reward 115/162 = 71%.
Figure 2. The Daily Percentages of Intervals On-Task for All Seven Children Per Condition.

Figure a. The daily percentage of intervals on-task for all seven children.

Figure b. The total percentage of intervals on-task for all seven children.

B₁: Baseline  
I₁: Visual Cueing  
I₂: Auditory Cueing and Self-Reward  
B₂: Baseline  
I₃: Visual and Auditory Cueing with Self-Reward  
I₄: Visual and Auditory Cueing with Self-Reward
Figure 3. The Daily Frequency of Teacher Praise Per Condition.

Figure a. The daily frequency of teacher praise by condition.

The Daily Percentage of Intervals On-Task

Figure b. The total percentage of intervals on-task per condition for all seven children.

- B1: Baseline
- I1: Visual Cueing
- I2: Auditory Cueing and Self-Reward
- I3: Visual and Auditory Cueing with Self-Reward
- I4: Visual and Auditory Cueing with Self-Reward
- # Field trip

- Baseline 2
- Mean = 18
- Mean = 13
- Mean = 14
- Mean = 13
- Mean = 16
- Mean = 13
Statistical Analysis

Each child's percentages of intervals on-task, by condition, were analyzed for statistical significance utilizing a Friedman distribution-free test (Hollander and Wolfe, 1973). The analysis revealed a statistically significant difference among the six experimental conditions ($S^1 = 15.66$, $\alpha$ approximately equal to .009). Therefore, the null hypothesis that there were no statistically significant differences among the six conditions was rejected. An analysis of the differences among conditions was completed using distribution-free multiple comparisons based on Friedman rank sums. The multiple comparisons revealed that compared to baseline the following conditions obtained differences which were statistically significant, auditory cueing with self-reward ($R_u - R_v = 2.21$, $1 - \alpha = .0530$), and the final introduction of visual and auditory cueing with self-reward ($R_u - R_v = 3.42$, $1 - \alpha$) values appear in Table 6. The difference between baseline 2 and the final experimental condition were analyzed using the Wilcoxon distribution-free signed rank rest. The difference was not found to be statistically significant for any $\alpha$ less than .406 ($T^+ = 16.5$).

Discussion of Data

The purpose of this study was to investigate the effects of visual cueing, auditory cueing and self-reward, and visual and auditory cueing with self-reward on the on-task behavior of learning disabled/behaviorally disordered children. The study was designed to gather evidence to answer five questions. The questions follow with the evidence gathered to answer the specific question.
Table 6

Values for 1-α Obtained in the Distribution-Free Multiple Comparisons Based on Friedman Rank Sums

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>1-α Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>I₁ - B₁</td>
<td>.4139</td>
</tr>
<tr>
<td>I₂ - B₁</td>
<td>.0530</td>
</tr>
<tr>
<td>I₃ - B₁</td>
<td>.0730</td>
</tr>
<tr>
<td>B₂ - B₁</td>
<td>.0095</td>
</tr>
<tr>
<td>I₄ - B₁</td>
<td>.0018</td>
</tr>
</tbody>
</table>

B₁ = Baseline; I₁ = Visual Cueing; I₂ = Auditory Cueing and Self-Reward; I₃ = Visual and Auditory Cueing and Self-Reward; B₂ = Baseline 2; I₄ = Visual and Auditory Cueing and Self-Reward.

1. Is the combined visual cueing and auditory cueing with self-reinforcement, effective in increasing the on-task behavior of students in a Learning Disabilities Behavior Disorders class?

All the children obtained percentages of observation intervals on-task, during the combined visual and auditory cueing with self-reward, which were higher than that obtained during baseline. The obtained increases ranged from 2% to 18% for individuals. The total differences for the group of seven children was a 9% increase: baseline, 67% on-task; visual and auditory cueing with self-reward, 76% on-task. However, when these procedures were removed, in congruence with
the reversal design, the measured on-task behavior increased for the group. The total percentage of on-task during Baseline 2 was 4% higher than the visual and auditory cueing with self-reward. The individual children's differences between the two procedures ranged from a decrease in Baseline 2 of 3% to an increase of 14%. When the visual and auditory cueing with self-reward were introduced after Baseline 2, another slight increase in on-task was obtained for the group: a 2% increase. The individual children's results ranged from a decrease of 13% to an increase of 12%.

Because no causal relationship was manifested between the visual and auditory cueing procedure with self-reward and on-task behavior, i.e., no decrease in on-task behavior was obtained during Baseline 2, the data are inadequate to answer question one.

2. Is the effect of visual cueing alone, different from the effect of auditory cueing with self-reinforcement? The individual results indicated that for most of the children the visual cueing was less effective than auditory cueing with self-reward in obtaining a higher percentage of on-task behavior. Auditory cueing with self-reward obtained percentages of on-task behavior ranging from a decrease of 3% to an increase of 10%. The total group percentage was 6% higher for auditory cueing with self-reward compared to visual cueing. Apparently, visual cueing was less effective than auditory cueing with self-reward during this study. When visual cueing alone was presented, a dramatic decrease in variability in on-task behavior was obtained. The decreased variability was not obtained when visual cueing was combined with auditory cueing.
3. Are the effects of the auditory cueing with self-reward and the auditory cueing with self-reinforcement different? The total percentages for intervals on-task for auditory cueing and self-reward, and visual and auditory cueing with self-reward were 75% and 76%, respectively. Auditory cueing with self-reward and the combined visual and auditory cueing with self-reward did not seem to be different in effectiveness in this study.

4. Is the student's self-evaluation of on-task behavior accurate: does the observer's evaluation match the child's self-evaluation? The results indicated that the children as a group self-rewarded, indicating they considered themselves on-task, and matched an observer's judgment, 71% of the events recorded. The percentage of agreement between the children as a group and the observer for accurate self-punishments was 42% of the events observed.

The data on accuracy of self-reward and self-punishment must be considered with caution due to the small number of observations: self-reward, 115/162 = 71% self-punishments, 10/24 = 42%. On the basis of the data obtained, self-reward appears to be accurate and self-punishment does not appear accurate by the children as a group. The individual accuracy of self-reward ranges from 55% to 77% agreement for the events observed. Self-punishment agreement ranged from 0, two children were not observed to self-punish, to 100% agreement. The raw data and percentages of agreement for self-reward and self-punishment by condition for each child are presented in Table 5.

5. Is the teacher's praise unaffected by the cueing procedure?
The daily frequencies of teacher praise and mean frequencies of teacher praise by condition are described in Figure 3. There did not appear to be a direct relationship between teacher praise and the cueing procedure. An increase in teacher praise was obtained during the first introduction of auditory cueing: an average increase of 5, from an average of 18 per experimental period. This average per condition was 13 and 14 during the next two conditions utilizing auditory cueing. These averages were only slightly above the average of 11 obtained during baseline. Therefore, no apparent relationship was obtained between the cueing procedures and teacher praise.

The results obtained from this study support the inference that due to an absence of any demonstrated functional relationship between visual and auditory cueing with self-reward and on-task behavior of the children in this study, no statement regarding the effectiveness of the experimental procedure can be made. The visual cueing procedure alone seemed to have little if any effect upon the magnitude of on-task behavior and the combined visual and auditory cueing with self-reward did not seem more effective than auditory cueing with self-reward. The student's self-evaluations of on-task behavior seemed accurate, but the self-evaluations of off-task behavior seemed inaccurate. Teacher praise did not seem to be functionally related to the cueing procedures. Statistical analysis of the data revealed that the differences among the experimental conditions were statistically significant. At a statistically significant level, auditory cueing with
self-reward, Baseline 2, and visual and auditory cueing with self-reward, i.e., the last experimental condition, obtained a higher percentage of intervals on-task, than baseline. The difference between Baseline 2 and the last experimental condition was not statistically significant.
CHAPTER V
SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Summary

This study evaluated the effect of visual cueing, auditory cueing with self-reward, and visual and auditory cueing with self-reward upon the on-task behavior of seven learning disabled/behaviorally disordered children. Also evaluated was the accuracy of the students' self-reinforcements and self-punishments. The frequency of teacher praise was measured to evaluate whether the cueing procedures were functionally related to the praise.

Six design conditions were implemented: baseline; visual cueing; auditory cueing with self-reward; visual and auditory cueing with self-reward; baseline 2; and, visual and auditory cueing with self-reward. The percentage of observed intervals the children were on-task and the frequency of teacher praise was measured during each session of the study. The accuracy of self-reinforcement and self-punishment were measured during each condition involving self-reinforcement.

Following baseline, visual cueing was introduced. Auditory cueing and self-reward were next. The next condition included visual and auditory cueing with self-reward. Baseline 2 followed with the final condition being the re-introduction of visual and auditory cueing with self-reward. The final intervention was selected because
it was the most effective in increasing the total percentage of intervals on-task for the children as a group. The re-introduction allowed the analysis of the effects within a reversal design.

The study was planned to gather evidence to answer the following questions:

1. Is the combined visual cueing and auditory cueing with self-reward, effective in increasing the on-task behavior of students in a learning disabilities class? The data were found to be inadequate to answer this question due to the failure to obtain a causal relationship between the independent and dependent variables.

2. Is the effect of visual cueing alone, different from the effect of auditory cueing with self-reward? A statistically significant difference was found with auditory cueing with self-reward more effective in increasing the percentage of intervals on-task.

3. Are the effects of the auditory cueing with self-reward, and the visual and auditory cueing with self-reward different? There was no statistically significant difference between the two procedures.

4. Is the students' self-reward for on-task behavior and self-punishment for off-task behavior accurate, i.e., does the observer's evaluation match the child's self-recording? Self-reward seemed accurate, self-punishment did not.

5. Is the teacher's praise unaffected by the cueing procedures? No functional relationship between teacher praise and the cueing procedures was discerned.
The study also gathered evidence to test the following null hypothesis:

1. There will be no statistically significant differences among the experimental conditions involving visual cueing, auditory cueing with self-reward, visual and auditory cueing with self-reward, baseline, and baseline 2.

The evaluation of the group data (Friedman distribution-free test) indicated that the experimental conditions were different at a statistically significant level (approximately equal to .009). A multiple comparison analysis (Friedman rank sums distribution-free multiple comparisons) revealed that intervention 2, intervention 3, baseline 2, and intervention 4 were statistically significantly different from baseline. No statistically significant difference was found between baseline 2 and intervention 4 or between intervention 2 and intervention 3.

**Discussion**

The results of the study are considered a stern test of the experimental procedures due to situational factors which did not increase the probability of obtaining favorable results. For example, the experimenter experienced difficulty in identifying an LD/BD class with a low percentage of intervals on-task. The identification process began in April. Though two different LD/BD teachers described their classes as difficult to control and very active, after a week of observation the experimenter discovered both classes to be on-task from 85% to 90% of the intervals observed. Apparently, most teachers
have increased on-task to high levels by April, the eighth month of the school year. The third class to be evaluated and the one selected for the study, did have a relatively low percentage of intervals on-task: 71%. This may, in part, have been related to the youngness of the group. Only one of the experimental subjects was of second grade age; the rest of the children were first grade age, i.e., six years of age before September 30th of the school year. The relatively low level of on-task behavior was not inferred as ineffective supervision or instruction because the teacher had sixteen years' experience as an LD/BD teacher and was considered very effective by the building principal and LD/BD supervisor. Therefore, the class was characterized as persistently off-task and young.

Another factor contributing to the conditions for obtaining unfavorable results was the teacher's stated dislike for the systematic use of tangible reinforcement. The teacher reported a belief that praise was a desirable tool, but the systematic use of specifically stated contingencies for activities or objects was not preferred. The teacher did agree to utilize the proposed reward system. The experimenter assured the teacher that each child would earn something each day there were rewards.

Due to the difficulty in identifying the experimental class, the study was not begun until March 1st. This resulted in the study being completed on the last day of school. The normal anticipation for the ending of school and the anticipated year-end parties did not enhance
the probability of obtaining favorable results during the last experimental condition.

The failure to obtain a reversal effect, fails to support an inference that the increased percentages of intervals on-task were functionally related to visual and auditory cueing with self-reward. Other inferences are highly probable. Perhaps the most parsimonious inference is that no functional relationship exists between the cueing procedures with self-reward and the percentage of intervals on-task. A large body of evidence evaluating self-reward has found a functional relationship between self-reward and other dependent variables, including on-task behavior (Bandura, 1969; Masters and Mokros, 1974). More specifically, a study used essentially the same procedures used in this study and did demonstrate a functional relationship between the auditory cueing with self-reward and on-task behavior (Glynn, Thomas and Shee, 1973). The samples were somewhat different. In the Glynn et al. study, the difficult to manage regular class children selected for the study were in their third year of school and about 5 months older.

Considering the research on self-reward and the known influence of the effect of self-reward, a possible inference is that weak reinforcement resulted in a weak effect. This may be true. Also, the relationship between the self-awarded x's and the rewards may have been weak. Due to scheduling difficulties, the period during which the reinforcers were obtained occurred about two hours after the experimental period. The experimental period ended about 11:50 A.M. and the reward period
occurred at 2:00 P.M. Several of the activities and objects were selected as reinforcers on the basis of the teacher's observations as to the popularity of the activities and objects with the children. During the reward period the children did seem excited about selecting reinforcers and did ask that some be repeated. On one occasion, the children earned a small box of crayons to keep. The first three children selected from two brands. When Subject 1 had a turn to choose, all of the more popular brand were gone and only the less popular was left. The child seemed very frustrated and cried. At the end of the study, the experimenter spoke with each child briefly and asked if they liked the "surprises." All reported liking the "surprises." When asked what their favorites had been, a wide range of objects and activities were named that roughly corresponded to what the teacher had described as the best probable reinforcers. The children's behavior during the reward period and during the post-experiment interview led the experimenter to infer that the reinforcers were probably effective, but a strong contingent relationship may not have existed between the self-recording in the morning and the presentation of reinforcers in the afternoon.

Statistically, there was a significant increase in on-task behavior with the introduction of auditory cueing with self-reward. But as Risley (1969) has indicated, a small increase in a desirable behavior which does not result in an acceptable rate of response is not significant in a socially meaningful sense. The 67% of intervals on-task during baseline was very close to the generally accepted level
for a classroom of 70% on-task. The increase of 9% to 76% on-task is not a meaningful improvement.

The fact that a gradual increase in on-task behavior did obtain during the study may be interpreted in one of three ways. One, maturation was primarily responsible. Two, as a result of the variable interval self-reinforcement schedule for so many sessions behavior resistant to extinction was learned, and no reversal was obtained when the experimental procedures were removed. Three, some other variable or variables were operating to obtain the described results. The resistance to extinction inference is difficult to assess. Glynn, Thomas and Shee (1973) measured on-task and preceded the final baseline with twenty sessions of class contingent reinforcement and then fifteen sessions of a self-control procedure, similar to the one used in this study. However, the final five sessions before the baseline were preceded by a four week absence of any experimental procedures. The Glynn and Thomas study (1974) which was very similar to this study, preceded the final baseline with ten sessions using the experimental procedure. In this study, baseline 2, the final baseline, was preceded by eighteen sessions utilizing self-reward, 80% more than Glynn and Thomas (1974). Well learned behavior, i.e., performance resulting from several sessions of variable interval self-reinforcement, can be highly resistant to extinction.

Since no strong inference can be made regarding the causality of the obtained behavior change, the basic question asked by the study, is intervention 3 effective in increasing on-task behavior,
must go unanswered. Other questions asked in the study did obtain relevant evidence to provide an answer.

Question 2, is the effect of visual cueing alone, different from the effect of auditory cueing with self-reinforcement, was answered. The statistically significant difference between the visual cueing alone and the auditory cueing with self-reward, indicated the visual cueing was less effective. This was as expected from the Glynn and Thomas (1974) study. Also, as Glynn and Thomas found, when visual cueing was first introduced, the variability of the measured on-task behavior was reduced. However, the second and third time visual cueing was introduced, this time in combination with auditory cueing and self-reward, no discernible effect on variability was obtained. One of the factors that might have contributed to the increased variability in later conditions involving visual cueing was the relatively small amount of apparent importance, teacher instructions for the group seemed to have. The routine for the reading period was well established and could be initiated with instructions taking less than sixty seconds. The experimenter asked the teacher to begin each experimental period with instructions and then give at least one group instruction during the period. This was usually contrived and the teacher instructions were usually short. In this study the visual cueing was not of much apparent benefit to the children. This apparent lack of importance of the visual cueing may well contribute to the lack of a difference between auditory cueing with self-reward and visual and auditory cueing with self-reward.
Previous studies measuring the accuracy of self-recording have obtained percentages of agreement between an observer and the child as low as 71% (Bolstad and Johnson, 1972) and as high as 79% after training (Fixen, Phillips and Wolfe, 1972). The results of this study were comparable to previous results. As a group, the children were accurate in self-rewarding 71% of the events observed. Also measured was the accuracy of the children's self-punishment, i.e., judging they were off-task when the auditory cue was emitted and then not recording an "x". No previous reports of accurate self-punishment have been published in the applied studies dealing with self-control. The children were accurate in their self-punishment 42% of the events observed. Some of the inaccurate self-punishment was observed when some of the children did not stop their work and self-reward, consequently inaccurately self-punishing themselves. There was no way to evaluate whether other self-punishments occurred because the children purposely deprived themselves of an "x" or whether they did not perceive the auditory cue. There is a possibility that some children had auditory discrimination difficulties. Occasionally a child asked the teacher if the auditory cue had sounded. On the average the auditory cue was given twenty-three times daily. The five children who obtained the highest percentages of accurate self-reward, self-rewarded less than the average number of occasions possible. The two children who had the lowest percentages of accurate self-reward, on the average, self-rewarded at least as often as the number of auditory cues. This means that though the average number of auditory cues was
23 times daily, the three children who were the least accurate in self-rewarding, usually self-rewarded a minimum of 23 times. Table 7 presents the total percentage of intervals on-task when self-reward was used for the study, the total accuracy percentage for self-reward and the percentage of possible reinforcements self-awarded per subject.

Table 7
The Percentage of Intervals On-Task, Total Accuracy Percentage of Self-Reward and Percentage of Reinforcement Self-Awarded During Self-Reinforcement Per Individual.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Percentage of Intervals On-Task With Self-Reward</th>
<th>Total Accuracy Percentage for Self-Reward</th>
<th>Percentage of Reinforcement Self-Awarded</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>78%</td>
<td>77%</td>
<td>86%</td>
</tr>
<tr>
<td>2</td>
<td>65%</td>
<td>55%</td>
<td>97%</td>
</tr>
<tr>
<td>3</td>
<td>89%</td>
<td>74%</td>
<td>92%</td>
</tr>
<tr>
<td>4</td>
<td>81%</td>
<td>67%</td>
<td>100%</td>
</tr>
<tr>
<td>5</td>
<td>73%</td>
<td>70%</td>
<td>99%</td>
</tr>
<tr>
<td>6</td>
<td>78%</td>
<td>69%</td>
<td>101%</td>
</tr>
<tr>
<td>7</td>
<td>67%</td>
<td>77%</td>
<td>97%</td>
</tr>
</tbody>
</table>

Glynn, Thomas and Shee (1973) reported the children tended to self-reward less than indicated by their mean percentage of on-task behavior. The obtained results do not corroborate those of Glynn et al. as all the children self-rewarded a higher percentage of the
possible reinforcement compared to the total percentage of intervals they were observed on-task. The Glynn et al. data are not exactly comparable to the data from this study because mean percentages were reported. Total percentages were used in this study. The experimenter questions the use of mean percentages, since percentages cannot be averaged unless the divisors are equal. For example, if the accuracy percentages for accurate self-reward in intervention 4 by Subject 1 were averaged they would be 57%. \( \frac{1}{2} \) 50\% \( \frac{3}{3} \) 100\% + \( \frac{4}{4} \) 100\% + \( \frac{1}{3} \) 33\% + 0/0 = 28\% + 5 or 57\%. The total percentage of accuracy, however, is 75\%: \( \frac{1}{2} + \frac{3}{3} + \frac{4}{4} + \frac{1}{3} + 0/0 \) = \( \frac{9}{12} \) = 75\%. Glynn, Thomas (1974) and Glynn, Thomas and Shee (1973) have used mean percentages in reporting on-task behavior also, which can be misleading and also make comparisons of results with this study difficult.

In attempting to compare this study with those of Glynn et al., and Glynn and Thomas other difficulties were encountered. The categorization of self-recording behavior became difficult. As the study progressed the children frequently spent time counting their x’s and on occasion took as long as 30 seconds to record an x after an auditory cue. On the average, 23 of 180 observation intervals or 13\% of the intervals involved some self-recording. If self-recording is considered on-task, the experimenter has increased on-task behavior as an artifact of the procedure. If self-recording is considered off-task, the experimenter has cued and indeed taught the children an off-task behavior. Self-recording was considered on-task
in this study if it occurred within thirty seconds of an auditory cue. Otherwise, self-recording was considered off-task behavior. Children who played with their card or counted the x's were considered off-task. No exact comparison of this study with the Glynn studies can be made because no distinction was made whether self-recording was on-task or off-task behavior. There is no doubt about the validity of the findings of the Glynn studies, but the comparison of measures is difficult with self-recording unspecified.

The experimenter attempted to measure the intervals of on-task as did Glynn and Thomas (1974), i.e., "... the child had to be observed in on-task behavior for the majority of the ten-second interval," p. 301. The experimenter was unable to obtain reliable measures using this definition. In effect a ten-second interval was not being used, a six-second interval was being used. This was too short an interval for the experimenter and the reliability observer to obtain acceptable agreement percentages.

It should be noted that the reliability agreement percentages were not reported as Hawkins and Dotson (1973) recommended. The authors recommended that when reporting scored interval and unscored interval percentages of agreement that the two percentages be described and then a mean percentage be computed. As previously discussed, percentages cannot be averaged unless the divisors, i.e., the denominators in the ratio, are equal. Therefore, the total percentage of agreement was presented. The experimenter found that the total percentage of agreement can be misleadingly high when either of the scored or unscored interval agreement is low and based on a small
number of observations. For example, a reliability measure for teacher praise was obtained during the forty-first session. The scored interval agreement raw data was 2/14 or 14% agreement. Since teacher praise happened infrequently most of the intervals were un-scored. The unscored interval agreement raw data were 166/178 or 93%. When the total raw data, 2/14 + 166/178 = 168/192, were converted to a percentage, an acceptable measure was obtained, 87% total agreement. Though arithmetically invalid, the averaging of percentages is a more stringent criterion for data reliability standards.

Recommendations

The relationship between the independent variables visual cueing, auditory cueing with self-reward and visual and auditory cueing with self-reward and the dependent variable, percentage of intervals on-task, may be better analyzed through the use of a multiple-baseline design. This would allow a lengthy treatment with no need to obtain a reversal effect. Also, the variables could be better assessed if the identified class or group of subjects obtained on-task measures near the 50% of intervals percentage, thus allowing for adequate improvement or a clearly described decrement. It is recommended that the reward period be scheduled immediately after the experimental period to maximize the incentive effects of the reinforcers. The visual cueing could be better assessed, if the experimental task would require teacher directions to the group as in most arithmetic or social studies lessons.
Due to the importance of improving self-control instruction with learning disabled/behaviorally disordered children, it is recommended the further work to analyze self-control and develop effective training be initiated with these children.

Some of the implications that can be drawn from the present study pertain to the apparent weak effect of the cueing procedures with self-reward upon on-task behavior. The use of a variable interval schedule of reinforcement in this study may not be conducive to rapid acquisition. Further research which might utilize a similar sequence of experimental interventions as this study, but evaluate the training effect of beginning with fixed interval cueing for reinforcement, e.g., fixed intervals of thirty seconds gradually lengthened to 120 seconds over a period of five days and then substituting a variable interval schedule. The evaluation of different schedules for self-reward during acquisition should prove helpful.

Also, for those children who produce the lowest percentages of accurate self-reward and self-punishment and for those children who seem to have auditory discrimination problems reducing their accuracy in responding to the auditory cue, the evaluation of the effect of cognitive modeling may prove helpful. Meichenbaum and his associates (e.g., Meichenbaum and Goodman, 1969; 1971) have investigated teaching impulsive and immature children, strategies in problem solving by modeling successful strategies and then rewarding the children for successful imitation. This procedure may be applicable in teaching LD/BD children to be more accurate in their self-rewarding and
self-punishment. The cognitive modeling may be an effective procedure in teaching appropriate on-task behavior also.
APPENDIX A

Self-Recording Form
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>
APPENDIX B

Self-Recording Form
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>9</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C

Recording Form
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>O-T</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
<tr>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
<td>OFF</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


Homme, L. E. Perspectives in psychology, XXIV: Control of coverants, the operants of the mind. *Psychological Record*, 1965, 15, 501-511.


Meichenbaum, D. Theoretical and treatment implications of development research on verbal control of behavior, Book in preparation, 1975.


