LOGAN, Mary Miller, 1942-
THE EFFECTS OF TYPE OF REINFORCEMENT AND SOURCE OF REINFORCER SELECTION ON INSTITUTIONALIZED MENTAL RETARDATES.

University of Arizona, Ph.D., 1970
Psychology, clinical

University Microfilms, Inc., Ann Arbor, Michigan
THE EFFECTS OF TYPE OF REINFORCEMENT
AND SOURCE OF REINFORCER SELECTION
ON INSTITUTIONALIZED MENTAL RETARDATE

by
Mary Miller Logan

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF PSYCHOLOGY
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

1970
I hereby recommend that this dissertation prepared under my direction by Mary Miller Logan entitled The Effects of Type of Reinforcement and Source of Reinforcer Selection on Institutionalized Mental Retardates be accepted as fulfilling the dissertation requirement of the degree of Doctor of Philosophy.

[Signature]
Dissertation Director

Date

After inspection of the dissertation, the following members of the Final Examination Committee concur in its approval and recommend its acceptance:

[Signatures]

[Signatures]

[Signatures]

Date

Date

Date

*This approval and acceptance is contingent on the candidate's adequate performance and defense of this dissertation at the final oral examination. The inclusion of this sheet bound into the library copy of the dissertation is evidence of satisfactory performance at the final examination.*
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Mary Helen Logan
IN MEMORY OF
ILYA
ACKNOWLEDGMENTS

The author wishes to thank the many people who contributed to the completion of this dissertation.

The author is especially grateful to Vincent J. Tempone, Ph.D., for his helpful direction in the formulation of the research design and in the writing of this dissertation and for the extra effort required on his part to direct the dissertation in absentia. Appreciation is also expressed to the remaining members of my committee--Drs. Ralph J. Wetzel, William B. Moore, Ronald H. Pool, and John L. Delk. Special thanks go to Mr. Lou Pfautsch, who designed and built the apparatus.

The author is grateful to the Superintendent, Mr. E. W. Killian, and the Research Committee of the Denton State School, Denton, Texas, for permission to use their facilities and residents in this research. Special thanks go to Mrs. Barbara Moss for the use of her office and testing suite and to Mr. Matt Moazami for his assistance with the subjects and in coordinating the research.

Appreciation is expressed to the Diagnostic and Evaluation Center of the Department of Pediatrics, University of Texas Southwestern Medical School at Dallas, for granting the leave of absence which made completion of this research possible.
Special personal thanks go to my husband, Daniel, and to Jeanne and Gerry Cain for their patience, understanding, and love, without which this dissertation could not have been completed.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>ix</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>x</td>
</tr>
<tr>
<td>1. BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>Summary</td>
<td>30</td>
</tr>
<tr>
<td>2. STATEMENT OF PROBLEM</td>
<td>32</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>32</td>
</tr>
<tr>
<td>3. METHODOLOGY</td>
<td>35</td>
</tr>
<tr>
<td>Subjects</td>
<td>35</td>
</tr>
<tr>
<td>Eligible Subject Pool</td>
<td>35</td>
</tr>
<tr>
<td>Subject Selection</td>
<td>37</td>
</tr>
<tr>
<td>Apparatus</td>
<td>38</td>
</tr>
<tr>
<td>Pilot Study A: Rankings of Reinforcement Values</td>
<td>42</td>
</tr>
<tr>
<td>Procedure</td>
<td>42</td>
</tr>
<tr>
<td>Results</td>
<td>43</td>
</tr>
<tr>
<td>Pilot Study B: Change Option in Setting of Performance Criterion</td>
<td>45</td>
</tr>
<tr>
<td>Procedure</td>
<td>46</td>
</tr>
<tr>
<td>Results</td>
<td>46</td>
</tr>
<tr>
<td>Experimental Procedure</td>
<td>48</td>
</tr>
<tr>
<td>4. RESULTS</td>
<td>53</td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>72</td>
</tr>
<tr>
<td>APPENDIX A: CA AND IQ DATA ON ELIGIBLE SUBJECT POOL AND EXPERIMENTAL SS</td>
<td>103</td>
</tr>
<tr>
<td>APPENDIX B: APPARATUS DIAGRAMS</td>
<td>106</td>
</tr>
<tr>
<td>APPENDIX C: RESULTS OF PILOT STUDY A</td>
<td>109</td>
</tr>
<tr>
<td>TABLE OF CONTENTS--Continued</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>APPENDIX D: PERFORMANCE RECORDING SHEETS AND EXPERIMENTAL DEMAND QUESTIONS</td>
<td>110</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>117</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                      Page
1. Mean Number of Crank Revolutions in     53
    Experimental and Control Groups        
2. Complete 2 X 2 X 2 Analysis of Variance 55
    of Performance Scores (Number of Crank
    Revolutions) of Experimental Groups
3. Comparisons of Control Group Means (Male 58
    and Female) with Means of Experimental
    Groups                                 
4. CA Distributions according to Sex of the 103
    116 Eligible Ss                        
5. IQ Distributions according to Sex of the 104
    116 Eligible Ss                        
6. Mean CAs according to Sex of Experimental 105
    Groups                                
7. Mean IQs according to Sex of Experimental 105
    Groups                                
8. Mean Reinforcer Rankings for Five Male 109
    and Five Female Pilot Study Ss
    Combined                              

viii
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Front Panel of the Apparatus</td>
<td>107</td>
</tr>
<tr>
<td>2. Experimenter's Remote Control Panel</td>
<td>108</td>
</tr>
</tbody>
</table>
ABSTRACT

The purpose of the current study was to investigate the effects of type of reinforcement, source of reinforcer selection, and sex on the performance of institutionalized mental retardates. It was hypothesized that external reinforcement and subject-selection of reinforcers would result in higher performance scores than self-reinforcement and experimenter-selection of reinforcer, that males would have higher performance scores than females, that Ss in the self-reinforcement conditions would set less stringent performance criteria but would self-administer greater magnitudes of token reinforcement than did the normal children studied by Bandura and Perloff in 1967, and that performance scores of the noncontingent reinforcement control group would be lower than those of the experimental groups.

Ss were 70 institutionalized mental retardates ranging in IQ from 50 to 70 and in CA from 12-8 to 34-6. All groups, consisting of seven male and seven female Ss, were well matched for mean IQ and mean CA.

The apparatus consisted of a crank which the Ss turned, a criterion-selector switch which could be set to performance criteria of 5, 10, 15, and 20, a vertical score-indicator panel, and a token dispenser button and chute.
This apparatus was connected to a remote panel located in an adjacent observation room, so that all functions of the apparatus could be remotely controlled by E.

Performance criteria and magnitudes of token reinforcement in the external reinforcement conditions were yoked to those variables set by Ss in the self-reinforcement conditions. In subject-selection of reinforcer conditions Ss worked for their choice of 26 items, while Ss in the experimenter-selected reinforcer conditions worked for a reinforcer specified by E. Control group Ss received only noncontingent reinforcement.

A 2 X 2 X 2 analysis of variance indicated that there were no significant main effects of type of reinforcement, source of reinforcer selection, or sex. Nor were there any significant differences between experimental group means. One of the causes of the high within-group variance which contributed to the lack of significant main effects was the significantly higher performance scores of twelve Ss who had contact with E subsequent to the instruction period.

In spite of the nonsignificant difference between external reinforcement and self-reinforcement, there was conclusive evidence that institutionalized retardates do not have well established processes of self-reinforcement; Ss in the self-reinforcement conditions made numerous
random changes in the setting of the criterion switch during the experimental procedure, indicating that for these Ss setting the criterion switch was not functionally equivalent to selecting a standard of performance. As hypothesized, Ss self-administered a significantly greater magnitude of token reinforcement than did the normal Ss studied by Bandura and Perloff. It was concluded, however, that in this population self-reinforcement was not contingent upon performance.

With sexes combined, there was no difference in performance between the control group and any of the experimental groups, but there was an unexpected significant difference between males and females in the control group. Separate comparisons of control group females and males with the experimental groups showed that control group females had significantly lower performance scores than all experimental groups, as hypothesized. However, there were no significant differences in performance between control group males and experimental groups.

Ss' responses to Experimental Demand Questions asked subsequent to the experimental procedure indicated that they were unable to conceptualize and/or verbalize the purpose of the study, the reasons they responded as they did, or the self-reinforcement process.
Results were discussed and speculations made regarding negative and unexpected findings. Suggestions for further research and for training institutionalized retardates in the process of self-reinforcement were offered.
CHAPTER 1

BACKGROUND

Reinforcement of human operant behavior is of two types, external and self-reinforcement. The extent to which behavior is controlled by its consequences depends upon the reinforcement values of the various stimuli that follow the behavior. In an institutional setting the staff should have excellent control over the behavior of the residents, provided that they have and make use of an adequate understanding of reinforcement principles and of the reinforcers that are effective with that particular population. However, research has not yet provided all of this information.

The current study sought to determine the effects of three independent variables on performance of institutionalized mental retardates: (a) type of reinforcement, self-reinforcement versus external reinforcement, (b) source of reinforcer selection, subject-selection versus experimenter-selection, and (c) sex of Ss. As the variables researched in the current study and the hypotheses made were determined by the results of previous studies and by the lack of studies in certain important areas, a review of
the relevant literature will be presented before the current study is discussed further.

In the majority of experimental studies with human Ss which have employed any type of reinforcer, the nature and magnitude of the reinforcers have been determined by the experimenter (E). These studies will be reviewed along with the few studies in which Ss have selected their own reinforcers.

The effects of different types or magnitudes of reinforcement depend on many variables, including the type of S, the nature of the experimental task, and the dependent variables under consideration. Studies using normal preschool and elementary school children will be reviewed first, beginning with those which compared the effectiveness of different types of reinforcers.

Terrell and Kennedy (1957) compared the effectiveness of verbal praise, reproof, candy, token, and light reinforcement on learning on a discrimination task. In acquisition the group receiving candy reinforcement was significantly superior to all other groups; in transposition, the candy reinforcement group was superior to all other groups except the token reinforcement group. On a maze-tracing task, Hanlon (1960) found no differences in performance of either nursery or elementary school children attributable to type of reinforcement (trinkets versus cereal), even though "earlier results indicated that
comparable Ss showed a highly significant preference for trinkets over cereal" (p. 267). Screven (1959) also used pilot studies, in this case to determine which type of reinforcer to use in subsequent studies. He reported that the "pilot studies involving candy vs. marble reinforcement did not indicate that candy was more effective than the marble tokens" (p. 464). Screven then studied the effects of marble reinforcement on performance, limiting the magnitudes of reinforcement to reinforcement versus non-reinforcement, conditions which produced significant differences in the running times of children pulling heavy weights down a runway.

Other studies with normal children have introduced greater variations in magnitudes of the same kind of reinforcer. With preschool children and a discrimination learning task, Weinstock (1965) found no differences in response speeds between the two magnitude of reinforcement conditions of one and five pieces of candy. On a different type of discrimination task, Miller and Estes (1961) found no significant difference in number of errors between groups receiving 50¢ and 1¢ as reinforcement, but both of these monetary reinforcement groups were inferior to the group receiving only knowledge of results.

Studies by Nakamura and associates (Nakamura and Ellis, 1964; Nakamura and Lowenkron, 1964) kept absolute magnitude of M & M or marble reinforcement the same but
varied relative magnitude of reinforcement for their high- and low-reinforcement conditions. A significant reinforcement effect was found with kindergarten Ss, but the data on third-graders yielded no such significant effect. Rather, the results emphasized the fact that the effect of different conditions of reinforcement may depend upon the interaction with other experimental variables, in this case Ss' level of persistence as rated by their teachers and parents.

A number of studies of magnitude of reinforcement effects have been conducted with the same type of apparatus, one in which the S's task is to pull a lever horizontally and the dependent variables are starting speed and speed of movement. In Bruning's (1964) study with kindergarten Ss, one and five pieces of candy per reinforced trial were the two magnitudes of reinforcement used. The effect of this variable was described as follows: "Larger magnitude of reward was found to have a slight, but nonsignificant decremental effect on performance" (p. 281). Bruning later (1965) studied the effects of a subsequent shift in magnitude of reinforcement. Contrary to prediction, Ss who received the smaller amount of reinforcement performed significantly faster than Ss receiving the larger amount of reinforcement. When magnitude of reinforcement was switched, "the mean speeds of Ss changed from large to small reward increased while the response level of Ss changed from small to large reward decreased" (p. 278).
The inferiority in performance produced by the larger magnitude condition was partially attributed to a satiation effect.

The results of the preceding studies on the effect of magnitude of reinforcement support Ryan's statement that "magnitude of reward has not proven to be such an effective variable in child studies" (1963, p. 11). He attributed the lack of significant differences between various magnitudes of the same material reinforcer to S's desire to please overpowering the effect of other reinforcers and to the lack of food deprivation in studies using candy as a reinforcer. In contrast to quantity or magnitude of reward, Ryan emphasized the potential importance of quality or value of reinforcement, proposing that high and low incentive groups should be composed by "noting each S's personal preferences in ranking a number of possible goal objects for their desirability" (p. 12). Using such a procedure Ryan had kindergarten children rank their preferences for six toys, and experimental groups worked for marbles which functioned as tokens which could be exchanged for backup reinforcers (highest or lowest ranked toys). However, the results indicated that neither the main effect nor any interactions involving incentive value were significant. Two similar unpublished studies (Estes, 1965; Bravenec and Estes, 1969) have yielded comparable findings; in both of these studies Ss (kindergarten and first-grade
children) worked for their most preferred toy as opposed to either least preferred toy or no reinforcement and received marbles as tokens on a continuous reinforcement schedule. In neither study was the reinforcement effect significant or involved in any significant interactions. Speculations as to why such negative results were obtained have included the following: (a) Ss' inherent interest in the task which "may have rendered any additional incentive motivation ineffective" (Ryan, 1963, p. 40), (b) lack of sufficient preference differential between high- and low-ranked reinforcers, and (c) the reinforcing effect of the marbles. As Ryan has stated, it is conceivable that "in the course of the 'game,' S became primarily concerned with the possibility of winning marbles and tended to forget about the toy" (p. 40). Certainly Screven's (1959) conclusions regarding the relatively high reinforcement value of marbles would support the validity of the latter speculation. In spite of his negative findings, Ryan reached the following conclusions with regard to the effects of incentive: "Whereas some incentive is generally more effective than none, greater amounts have not, as yet, produced faster learning; certain incentives facilitate learning more than others and perhaps the optimal incentive condition is that which allows S to choose his own reward" (p. 76). Bruning has concurred, with his statement that "while performance
tends to be positively related to quality of reward... an inverse relation appears to be true when quantity is varied" (1965, p. 281).

Several writers have emphasized the importance of research on reinforcement variables with mental retardates (Heber, 1959; Watson, 1967; Watson and Lawson, 1966; Watson, Orser, and Sanders, 1968). As Heber has stated, "a problem of great importance in training and habilitation is the question of what incentive conditions, and variations in incentive conditions mental retardates of various levels will respond to" (1959, p. 667). Watson has stressed the need to "develop a variety of reinforcing events, both positive and negative, and systematically evaluate their relative effectiveness. There is a great need for more powerful reinforcement procedures. The generalized or token reinforcement may be potentially the most effective and durable reinforcer of all" (1967, p. 14). However, it is necessary to have effective primary reinforcers in order to develop effective generalized reinforcers. Watson and Lawson (1966) concluded that researchers have "only begun to discover truly effective, long-lasting reinforcers. Until these are better developed, certain important long-term learning research is going to be seriously hampered" (p. 16).

Results of studies of reinforcement effects with mental retardates have in many ways paralleled the above
findings with normal children. However, there have been more studies with mental retardates than with normals comparing the effectiveness of different types of reinforcement. Andronico (1964) compared the effectiveness of food, social, visual, and no reinforcement with mental retardates and found that food and social reinforcement produced higher response rates than did visual reinforcement or the control condition. Boys had a higher response rate than girls, and there was also a sex difference in reinforcement effects, with retarded girls' highest response rate under the social reinforcement condition and retarded boys' highest response rate for food reinforcement.

Black and Thomas (1966) studied the relative effectiveness of primary (M & M), secondary (token), and social reinforcement on a discrimination task with institutionalized mongoloids. The group receiving token reinforcement performed significantly better than the group receiving candy reinforcement, and the social reinforcement group did not differ significantly from the other two. In attempting to account for the finding of superiority of secondary over primary reinforcement, the investigators considered the possibilities that the candy reinforcement distracted the Ss from the task and that the tokens or the act of redeeming the tokens "involves some intrinsic reinforcing properties" (p. 586). Haywood and Weaver (1965) compared the effects of four reinforcement conditions on the speed of hole
punching of institutionalized retardates and found that the high monetary group (10¢ reward) was inferior in performance to groups receiving low monetary reward (1¢), task motivation (the opportunity to perform another task as reinforcement), and no reinforcement. There were no significant differences among these three reinforcement conditions. In a study of noninstitutionalized mental retardates' performance on four different learning tasks, White (1968) found no significant differences between any of the reinforcement conditions (social, candy, buzzer, and control). The lack of differences in the reinforcement conditions was attributed to the effects of subject selection and the novelty of the teaching apparatus and the tasks.

Watson and his associates (Watson, 1967; Watson, et al., 1968) have made attempts to determine what reinforcers are effective with mental retardates. In the study by Watson, et al. (1968), severely mentally retarded Ss pulled a lever to obtain chips which could be exchanged for a variety of reinforcers, including candy, toys, music, and films. Initially more chips were spent for manipulable items, but a gradual preference for edibles emerged. There was no overall significant difference between preference for edibles versus non-edibles. Music was significantly preferred to toys, but there were no overall significant preference differences for different kinds of candy.
Wolfensberger (1960) compared the effects of "concrete" and "symbolic" reward and punishment (presentation or withdrawal of a material reinforcer or token). No significant main effects or interactions were found for these reinforcement variables. Baumeister and Ward (1967) found money and praise to be "about equally efficacious" in a reaction time task with institutionalized male retardates in which reinforcement was contingent upon speed of response. These investigators attributed their positive results to this contingency and contrasted their design with that of Wolfensberger (1960) in which reinforcement was dispensed on a fixed ratio schedule and was not contingent upon speed of performance.

Different magnitudes of the same reinforcer have been used to define "high" and "low" reinforcement conditions in a few studies with mental retardates. In general, such conditions have failed to yield the predicted differences in performance. Cantor and Hottel (1955) found no significant differences in performance on a simple discrimination task by their mentally retarded Ss as a function of their differential reinforcement conditions of small and larger amounts of peanuts. Likewise with mental retardates, Cromwell and Moss (1959) found that while expectancy statements were a function of anticipated reinforcement, there was no significant difference between their "low reward condition" of 5¢ worth of candy or gum and their "high reward
condition" of 25¢ worth of candy, gum, or cigarettes. In Ellis' (1962) study, institutionalized male retardates pulled a manipulandum on an increasing fixed ratio schedule for differential magnitudes of reinforcement of one or three cigarettes. Although the difference was not significant, the trend was for the group receiving the smaller magnitude of reinforcement to emit a higher rate of response, to make more responses to extinction, and to remain in the situation longer. There were no significant differences in two conditions of reinforcement magnitude in a second experiment in which Ss chose to work for either candy or cigarettes. Kahn and Burdett (1967) likewise found no significant performance differences according to their reinforcement conditions in which mentally retarded Ss received a fixed reward of three pieces of candy at the end of the session, one piece of candy per trial regardless of performance, or one piece of candy on every trial in which performance improved over that in the previous trial.

In contrast to the negligible effects of different magnitudes or quantities of reinforcement are the findings of studies in which Ss in some way have determined their own reinforcers according to their qualitative value. As Bijou and Sturges (1959) have pointed out, stimuli referred to as "rewards" or "incentives" may not function as such for a particular S. They only way it can be determined whether a stimulus functions as a reinforcer is by
empirical means: "Observe the behavior prior to the introduction of the stimulus and note the changes in behavior after presentation contingent upon a response. If the frequency of response is increased, then the stimulus is a reinforcer for that organism in that situation" (p. 151). Obviously employing such a procedure for determining the reinforcing properties of each stimulus for each S in each situation would be tedious and time-consuming. Therefore, other methods of determining Ss' reinforcement values have been developed. In the paired-comparisons method, an S is presented with a series of pairs of stimuli and indicates which of the two he prefers. In the original study employing this method Witryol and Fischer (1960) presented five objects in pairs to normal preschool children. The overall order of preference for the objects was bubble gum, balloons, charms, marbles, and paper clips, and the results indicated that "although our preschool Ss could make reliable decisions on these paired incentive objects, there were strong individual differences for final scale values" (p. 472). Tyrrell, Witryol, and Silverg (1963) used the same procedure and stimuli with a group of institutionalized mental retardates and found high reliability and high inter-subject agreement in preferences. The incentive ranks obtained with this population were identical, with only one exception, to those obtained by Witryol and Fischer (1960) with normal preschool children.
Four studies have related reinforcer preferences established by the paired comparisons method to measures of performance. Blount (1967) found incentive value to be a function of both qualitative and quantitative variables; highly preferred items included a billfold, a large bag of M & Ms, and a package of cigarettes, while the items ranked as having the lowest values were one M & M and one cigarette. A significant reinforcement effect was found on performance on a finger maze; the high-reward group (working for a billfold) was superior in performance to the low-reward group (working for one M & M). Siegel, Forman, and Williams (1967) used a method of paired comparisons in which preferences were determined by the number of responses made by their mentally retarded Ss working for each of the two items. This method yielded information as to what reinforcers an S would work for rather than relying on verbal statements of preference. Ss' preferences for the objects were as follows: penny, Tootsie Roll, candy corn or chiclet, washer, and nothing. Ss then performed a discrimination task in which one of each pair of stimuli was consistently reinforced by a highly-valued reinforcer (Tootsie Roll) and the other was consistently reinforced by a low-preference reinforcer (washer). Ss consistently selected the stimuli associated with the highly-valued reinforcer. In a similar study, Siegel, Williams, and Forman (1967)
used a penny and a washer to define their high and low reinforcement conditions. Groups were run at one level of reinforcement, then switched to the other; the group switched from the low to the high reinforcement condition exhibited a significant increase in response rate, but Ss switched from high to low reinforcement exhibited little change in response rate. On the basis of these findings, it was concluded that the response rate of mental retardates on a simple performance task was "sensitive to variation in the preference value of the incentive or reinforcer" (p. 453). Williams (1968) also used the penny and washer as high and low reinforcers in a study of reversal learning. There was a significant difference between the two reinforcement conditions, indicating that high reinforcer value facilitated reversal learning in noninstitutionalized mental retardates:

The second method by which Ss have indicated reinforcer preference was that used by Heber (1959), in which Ss were shown all available items to be used as reinforcers and were asked to indicate which item they would select if they had their choice of one of the items. When the S made a choice, that item was removed and the procedure was repeated, until all items were ranked in the order of the S's preferences. Heber used consumable and nonconsumable material items, and his Ss were institutionalized mental
retardates ranging in CA from 16 to 44 years. There was little inter-S consistency in the rankings. Two groups of Ss completed a complex formboard, working for high and low incentive conditions (Ss' first and seventeenth choice of reinforcers, respectively). Ss in the high incentive condition performed significantly faster than did Ss in the low incentive condition, and a reversal in incentive conditions resulted in a corresponding reversal in efficiency of performance for the two groups of Ss. Heber attributed these significant results to having had the Ss rank their own reinforcement values and made the following comments regarding use of such a technique (p. 671):

The low inter-individual consistency which was obtained in the preference ranking of reward objects raises a question as to the desirability of arbitrary selections of a uniform reward by the experimenter in studies of this type with mentally retarded subjects. The incentive scaling technique used in this study offers a method for at least roughly equating the value of the reinforcement for different subjects in learning studies with the mentally retarded.

On the basis of the findings of the above studies, it can be hypothesized that reinforcers selected by the S would have more effect upon his performance than would reinforcers specified by the E. If such a finding were empirically verified, it would have important implications not only for behavioral research but also for the reinforcement practices in institutions. Although institutions vary somewhat in this respect, an institution typically offers
few positive reinforcers and those reinforcers are determined by the staff rather than by the persons being reinforced and are presented on a noncontingent basis. If it were demonstrated that S-selected reinforcers are more effective in maintaining behavior of institutionalized mental retardates than are E-selected reinforcers, such a finding would represent a powerful argument for institutions' introduction of more choices of reinforcers on the part of their mentally retarded residents.

Three other issues concerning reinforcement should also be considered here: (a) the importance of instructions, (b) the interaction between reinforcement and other experimental variables, and (c) the contamination of the effect of reinforcement variables by uncontrolled social reinforcement.

Evans and Spradlin's (1966) study clearly demonstrated the important influence of instructions in reinforcement studies. These investigators used money as their reinforcer but varied the schedules of reinforcement, using a "piece-rate plan," in which a unit of pay was given for each unit of work so that total pay depended upon total amount of work, and a "salary plan," in which a specified amount of pay was presented at the end of a specified unit of time regardless of amount of production. Institutionalized retardates were significantly more productive in the piece-rate condition, but in absolute productivity there
was only about a 10% difference between the two groups. This led the investigators to question whether the instructions given to the salaried group ("You are being paid $1.00 to pull the knob as many times as you can") led Ss to believe that payment was contingent upon productivity. When these Ss were run on a no-salary condition ("I will not pay you any money to pull the knob; however, I want you to pull the knob as many times as you can"), the response was significantly less than under the salary condition. The investigators concluded: "Apparently money, even when it is noncontingent on responding, does lead to a higher response rate than when no money is involved" (p. 132). However, due to the high rate of response in the no-salary condition, it was concluded that instructions were "a potent antecedent variable with high level retarded Ss" (p. 132). Ss were then run on a no-salary condition and instructed not to pull the knob; this condition produced a mean number of responses of .80. The high response rate was then restored by instructing Ss to pull the knob as many times as they could. The effectiveness of instructions was considered to be due to Ss having been reinforced in the past for following instructions and punished for not following them. Headrick (1963) also emphasized the important influence of instructions on the response rates of mental retardates, to the extent that "the use of definitive verbal instructions ... introduces cues into the
conditioning situation such that subsequent reinforcement contingencies cease to be the primary factor controlling the emission of responses" (p. 431).

Haywood and Weaver's (1967) findings emphasized the important influence of other variables on the effect of reinforcement conditions. They used two groups of institutionalized mentally retarded Ss who were characterized as "relatively extrinsically-motivated" and "relatively intrinsically-motivated." The experimental task was hole punching and the incentive conditions were 10¢ reward, 1¢ reward, task incentive (promise of another task), and no reward (control). The results indicated that "IM subjects punched more holes under the task incentive than under money incentives, while EM subjects punched more holes under the money incentive conditions" (p. 459).

When social reinforcement is used in an experiment but is not an experimental variable, it can prove to be a contaminating factor. Cantor and Hottel (1955) considered the presence of social reinforcement in their study to be a confounding factor which tended to equalize the effects of the experimental reinforcement conditions and render them ineffective. Wolfensberger (1960) warned against use of social reinforcement in addition to material reinforcement, particularly with mentally retarded Ss, pointing out that differences in performance found in such studies "may well be due to E's saying 'you are right' or 'you are wrong'
rather than to the giving or taking, increasing or decreasing the number of prizes or chips. . . . to a mentally defective who had experienced failure and punishment all his life, being right means more than winning a candy bar" (p. 906). Bijou and Sturges (1959) have also warned against the confounding influence of social reinforcement and have urged that "reinforcers other than those experimentally introduced be maximally controlled by using minimum amount of instruction . . . delivered in a consistent and standard fashion and by presenting the manipulables in such a way that social factors are eliminated, controlled, and evaluated" (p. 168).

All of the studies reviewed above were studies of external reinforcement. Regardless of whether the reinforcers employed were selected by E or S, they were always dispensed by E contingent upon some predetermined behavior of S. An overwhelming majority of the reinforcement studies in the literature have used external reinforcement; it is only recently that self-reinforcement has been submitted to experimental analysis. The importance of self-reinforcement and the results of available studies of self-reinforcement will be considered next.

Skinner (1953) discussed self-reinforcement and self-control (pp. 237-238):

Self-reinforcement of operant behavior presupposes that the individual has it in his power to obtain
reinforcement but does not do so until a particular response has been emitted. The ultimate question is whether the consequence has any strengthening effect upon the behavior which precedes it.

It was not until the past decade, however, that experiments were conducted to examine the factors influencing the development of self-reinforcement and the role self-reinforcement plays in the maintenance of behavior.

Marston has elaborated upon Skinner's statements in defining self-reinforcement as "the delivery of a reinforcing stimulus by an organism to itself without direct and current external controls. The reinforcements are freely available but delivered under contingencies specified only by the organism that is both delivering and receiving them" (1964, p. 879). Bandura (1969) has stressed the importance of self-reinforcement in the following statements: "Although the controlling power of externally occurring consequences cannot be minimized, self-reinforcement may frequently outweigh the influence of external outcome in governing social behavior, particularly in the case of older children and adults" (p. 32), and "The establishment of self-monitoring reinforcement systems is essential if induced behavioral changes are to generalize and endure, particularly where social environments provide either weak support for new modes of behavior or conflicting patterns of reinforcement" (p. 624). Although the
conditions facilitating and determining the acquisition of behavioral standards and self-reinforcing responses have by no means been conclusively established, several of the determinants of self-reinforcement have been investigated experimentally.

Numerous studies (Allen and Liebert, 1969; Bandura, Grusec, and Menlove, 1967; Bandura and Kupers, 1964; Bandura and Whalen, 1966; Grusec, 1966; Herbert, Gelfand, and Hartmann, 1969; Kanfer, 1966; Liebert and Fernandez, 1970; Liebert, Hanratty, and Hill, 1969; Marston, 1965a; McMains and Liebert, 1968; Mischel and Liebert, 1966; Ora, 1969) have concerned themselves with the importance of models and vicarious reinforcement on the adoption of behavioral standards for self-reinforcement. These studies have indicated that models exert a significant influence on Ss' adoption of behavioral criteria and magnitude of self-reinforcement (Bandura and Kupers, 1964; Herbert, et al., 1969; Kanfer, 1966; Ora, 1969), that positive vicarious reinforcement enhances the model's effect (Bandura, et al., 1967; Liebert and Fernandez, 1970), that adults are more powerful models than peers (Bandura and Kupers, 1964), but that the presence of conflicting peer standards decreases an adult model's influence (Bandura, et al., 1967). Models' experience (Allen and Liebert, 1969) and competence in comparison with that of the Ss (Bandura and Whalen, 1966) are relevant variables, as is the degree of the model's nurturance (Bandura,
et al., 1967) or "rewardingness" (Grusec, 1966) The influence of a model is enhanced by direct reinforcement of the S for imitation (Liebert and Fernandez, 1970) and decreased by incentives for nonimitation (Allen and Liebert, 1969), discrepancies between the observed and imposed reinforcement criteria (Mischel and Liebert, 1966), and discrepancies in self-reinforcement criteria exhibited by two models (McMains and Liebert, 1968). The criteria used by a S in reinforcing another person are closely related to those used in self-reinforcement (Marston, 1965a; Mischel and Liebert, 1966).

Several studies have demonstrated that criteria and magnitude of self-reinforcement are significantly influenced by direct reinforcement, instructions, and other "training procedures." Kanfer and Marston (1963a, 1963b) have demonstrated that standards of self-reinforcement can be directly conditioned by means of social reinforcement and that self-reinforcement established in this manner transfers to other tasks, confirming their hypothesis that "pretraining can significantly affect a person's tendency to reinforcer his own behavior, when there is no external source of information about the accuracy of his responses" (1963a, p. 68). Mischel and Liebert noted that "in all treatment conditions subjects without exception rewarded themselves in the model's presence whenever she indicated
that the score was deserving and never when she commented negatively on the performance" (1966, p. 49). Kanfer and Duerfeldt (1967) found that Ss' self-evaluations of performance paralleled previous E-evaluations and that self-reinforcement was comparable to that of the training level, but that self-evaluations and E-evaluations of performance did not significantly affect rate of self-reinforcement.

Two studies have examined the effect of rule structure on self-reinforcement. Liebert and Allen (1967) reported that in keeping with their prediction, "high rule structuring resulted in less rule deviation and more self-rewarding verbalizations than did low rule structuring" (p. 445). Likewise, Liebert, et al. (1969) found that the higher the level of rule structure, the greater Ss' adherence to a stringent criterion of self-reinforcement. They concluded that their study provided "striking evidence of the importance of verbal communication for influencing children's adoption of self-imposed standards" (p. 100).

An important area of inquiry concerns the influence of past experiences of success and failure on self-reinforcement; this issue is especially relevant to the consideration of self-reinforcement in the mentally retarded. Available studies of the influence of past experiences of success and failure upon self-reinforcement (Bandura and Whalen, 1966; Masters, 1968, 1969; Mischel, Coates, and Raskoff, 1968) have yielded some apparent discrepancies,
with obtained results depending on the nature of the tasks involved, the contingent versus noncontingent nature of the self-reinforcement, and other experimental variables. Bandura and Whalen (1966) defined their antecedent success-failure conditions in terms of Ss' performance on a series of tasks being compared favorably or unfavorably with that of their partners and with fictitious normative data. However, contrary to prediction, there was no overall significant difference in the magnitude of self-reinforcement as a function of prior success or failure experiences, and trends in this variable were dependent upon treatment conditions and level of performance. Exposure to experiences of success resulted in significantly more self-reinforcement in children exposed to all three modeling conditions than in control groups, while prior experiences of failure decreased self-reinforcement among children exposed to an inferior model but increased it among children in the control group, the latter finding suggesting that "under certain circumstances self-gratification may primarily serve a therapeutic rather than a self-congratulatory function" (p. 381).

Mischel, et al. (1968) investigated the effect of success and failure on an achievement task on subsequent noncontingent self-gratification in a situation which was "objectively irrelevant to the initial success or failure
task (p. 381). Ss who were successful engaged in a greater amount of self-gratification than did Ss who had failed. Masters (1968) focused on the social comparison aspect of success-failure experiences and their effects upon self-reinforcement. Ss who had less success than their partners and girls who had been more successful than their partners were generous in subsequent noncontingent self-reinforcement, while Ss whose success was equivalent to that of their partners were more stringent in subsequent self-reinforcement. In a later study Masters (1969) found a strong tendency for Ss who received fewer tokens than their partners to indulge in subsequent high self-reinforcement. It is likely that the effects of long-term, cumulative experiences of success and failure on self-reinforcement are even more complex and difficult to determine than the effects of immediate, experimentally-induced success-failure experiences that have been reviewed.

Other studies have examined the determinants of accuracy of self-reinforcement, with accurate self-reinforcement defined as the administration of self-reinforcement when it is objectively warranted by the quality of the preceding response. Accuracy of self-reinforcement has been shown to be a function of degree of learning (Kanfer, Bradley, and Marston, 1962; Kanfer and Marston, 1963b); age, with a gradual decline of inappropriate self-reinforcement with increasing age (Kanfer, 1966);
incentive (Kanfer, 1966; Marston and Kanfer, 1963); and class standing, with low-ranking Ss taking more inappropriate self-reinforcements than high-ranking Ss (Kanfer, 1966). With college students as Ss, Marston (1964) found that the overall rate of self-reinforcement was increased by a reduction in ambiguity of the task. This finding would tend to indicate that such Ss evaluate accuracy of their responses and self-reinforce accordingly.

Although the relative influence and exact role played by each of the determinants of the development of self-reinforcement have not been clearly delineated at the present time, it can be concluded that the development of self-reinforcement standards and behaviors is influenced to a certain extent by all of the following factors: (a) observation and imitation of the behaviors and criteria for self-reinforcement exhibited by models, (b) differential reinforcement by external agents of behavior which meets with those agents' own criteria of acceptable or desirable behavior, (c) social comparison processes, in which the attainments of others are used as the norm by which a person evaluates his own behavior, and (d) experiences of success and failure.

As Skinner (1953) has stated, the "ultimate question" regarding self-reinforcement is whether it can maintain and strengthen behavior. The most important study investigating the question of whether self-administered
consequences actually possess reinforcing properties and can therefore maintain behavior has been that of Bandura and Perloff (1967). The results of this study indicated that self-reinforcement and external reinforcement were equally effective in maintaining behavior in normal children and that both of these types of reinforcement were significantly more effective than either noncontingent reinforcement or no reinforcement. An important finding in this study was that these Ss set stringent criteria for self-reinforcement, even though they were in a situation characterized by "high permissiveness for self-reward" (p. 115) and adoption of stringent standards decreased the amount of reinforcement obtained.

Two additional studies (Johnson, 1969; Kelly, 1965) have demonstrated the efficacy of use of self-reinforcement in obtaining desired effects on behavior and learning. Johnson described his findings with elementary school children: "Self-reinforcement maintained attention and discrimination behavior at the same high level as did external reinforcement with no decrement in discrimination accuracy . . . . The results indicated . . . that self-reinforcement procedures can produce and maintain behavior change" (1969, p. 4367). Kelly found self-reinforcement to be more effective than external or no reinforcement in paired-associates learning by schizophrenic women, possibly because in this population self-reinforcement "may have served to focus"
attention on the task at hand and thus to reduce preoccupation with task-irrelevant stimuli of various kinds" (1965, p. 2869).

Bandura and Walters have concluded that "most persons will acquire discriminative self-controlling behavior as a consequence of exposure to differential modeling cues and differential patterns of reinforcement" (1963, p. 222). One of the primary questions examined in the current study was whether this statement applied to institutionalized mental retardates, whether they could maintain their own behavior by means of self-reinforcement or whether there were deficiencies or differences in self-reinforcement standards and behaviors among this population.

In spite of the innumerable reinforcement studies that have been conducted with mental retardates and the rapidly growing interest in self-reinforcement, only one study has appeared in the literature which examined self-reinforcement in the mentally retarded. Brodsky, LePage, Quiring, and Zeller (1970) used the term "self-evaluative responses" or "SERs" rather than self-reinforcement, defining SERs as those responses "by which the individual matches other aspects of his behavior to the behavioral standards specified by the contingencies of reinforcement." They discussed the properties of SERs further (p. 792):
An accurate evaluation by an individual that his behavior has reached, or is reaching, the required criteria for reinforcement should serve as a discriminative stimulus, indicating the occurrence of reinforcement following sequence completion. Thus, SERs should function as do other response-produced stimuli in behavioral sequences and should acquire conditioned reinforcer properties.

These investigators studied the baseline level and modifiability of SERs of institutionalized retardates in a matching-to-sample task and reported a "tendency for the institutionalized male retarded adolescent to over-evaluate his performance" (p. 792). Subsequent experimenter-administered reinforcement of accurate responses resulted in a significant increase in accuracy of SERs in the experimental group without a decrease in total SER rate. Brodsky, et al. discussed remaining areas in which research needs to be done: "The exact role of SERs remains to be fully delineated. Further, it remains to be demonstrated that the establishment of accurate self-evaluations does, indeed, lead to a subsequent maintenance and generalization of behavior" (p. 795).

The results of the study by Brodsky, et al. (1970) had important implications for the current study. The first important finding was that institutionalized mental retardates tend to "over-evaluate" their performance and the second that such a tendency could be modified by contingent reinforcement of Ss' responses. The current study was designed to yield more information regarding
self-reinforcement in institutionalized mental retardates, including the important issues of whether self-reinforcement could maintain behavior in this population and whether self-reinforcement was as effective as external reinforcement.

**Summary**

From the studies reviewed it can be concluded that variations in the magnitude of the same material reinforcer have generally not produced significant differences in performance. One of the reasons for such negative findings could be that the particular reinforcers used, which were selected by E, were not reinforcing to the Ss. It has been hypothesized by several investigators that more significant performance differences could be obtained by varying the qualitative values of the reinforcers used, that is, defining high and low reinforcement conditions according to Ss' own rankings of preferences for the various reinforcers used. The results of studies using such an approach have varied, and it appears that the results may depend upon the range of preference differential in the items ranked, as well as upon various other features of the experimental design, including use of token reinforcement, whether Ss are used as their own controls, the confounding presence of social reinforcement, and the nature of the task and the dependent variables. While findings of previous studies
tend to indicate that varying reinforcer conditions along qualitative dimensions is more effective than varying them along quantitative dimensions, additional research needs to be done to determine under which conditions this approach is most effective.

Self-reinforcement has been shown to be greatly influenced by modeling procedures, direct and vicarious reinforcement, instructions, and previous experiences of success and failure. The few available studies indicated that with normal school children self-reinforcement was as effective as external reinforcement in maintaining behavior. However, the effectiveness of self-reinforcement in maintaining behavior in institutionalized retardates remained to be explored.
CHAPTER 2

STATEMENT OF PROBLEM

The current study was designed to explore some of the unanswered questions that were raised in the review of the literature. It was designed to evaluate with a population of institutionalized retardates the influence on performance of type of reinforcement, source of reinforcer selection, and sex. It provided direct comparisons of the effectiveness of self-reinforcement versus external reinforcement and of subject-selection of reinforcer versus experimenter-selection of reinforcer.

Hypotheses

The following hypotheses were made regarding the results of the current study.

1. There will be significantly higher performance (more crank-turning responses) in the subject-selected reinforcer groups than in the experimenter-selected reinforcer groups. Contributing to the effectiveness of the subject selection of reinforcers will be the fact that these Ss are working not only for a reinforcer but also for the opportunity to make a choice of reinforcer, which is
likely a relatively infrequent occurrence in institutions for the mentally retarded.

2. Due to the factor of institutionalization, self-reinforcement and external reinforcement will not be equally effective in maintaining behavior in this population, as Bandura and Perloff (1967) found them to be with normal elementary school children. Due to the lack of adequate models and of consistent direct reinforcement of behaviors meeting the reinforcing agents' criteria for acceptable behavior in an institution, it is hypothesized that these Ss will not have developed effective self-reinforcement standards and behaviors and that consequently self-reinforcement will be less effective than external reinforcement in maintaining behavior in institutionalized mental retardates.

3. The overall sex difference in performance attributed by Bandura and Perloff (1967) to greater strength and endurance on the part of the male Ss will also be found in the present study. Whereas Bandura and Perloff found an additional sex difference, that external reinforcement was more effective than self-reinforcement for boys only, it has been predicted (Hypothesis 2) that in the current
study such will be the case for all Ss, rather than there being a sex difference in the effects of this variable.

4. In comparison with the Ss in the Bandura and Perloff study, Ss in the self-reinforcement conditions in the current study will set less stringent performance criteria and will take a greater number of tokens/reinforcement period.

5. The noncontingent reinforcement control group will have the lowest performance of all groups, confirming the hypothesis that behavior is not maintained by noncontingent reinforcement. Such a finding would have important implications for changes in typical institutional practices, in which the relatively rare positive reinforcement that does occur is administered on a noncontingent basis.
CHAPTER 3

METHODOLOGY

Subjects

Subjects (Ss) used in both the primary experiment and the two pilot studies were mental retardates residing at Denton State School whose IQs ranged from 50 to 70 and whose MA were 7 years and above. The latter restriction was made in order the comparisons might be made between the performance of these Ss and that of the normal children aged 7-10 years who served as Ss in the Bandura and Perloff (1967) study. Ss in the current study ranged in CA from 12-8 to 34-6.

Eligible Subject Pool

The original subject pool from which the experimental and pilot study Ss were selected consisted of all residents of dormitories designated as housing the mildly retarded. There were a total of 227 potential Ss in this pool, 137 males and 90 females. From this pool Ss were eliminated whose IQs were too high (above 70) or too low (below 50), whose MA were too low (below 7 years), who had severe sensory deficits (blindness or deafness) or neuromuscular impairments, and who were characterized as being uncooperative, unresponsive, or psychotic.
Ss remaining consisted of 68 males and 48 females who met the criteria of an MA of at least 7 years, an IQ in the 50-70 range, and no sensory or neuromuscular impairments which would interfere with their meeting the demands of the experiment. As in the original pool of 137 male and 90 female Ss, the remaining pool of Ss had a sex ratio of roughly 3 males : 2 females. The male Ss ranged in CA from 11-5 to 34-6, and the female Ss from 12-8 to 33-11, but the CA distributions differed for the two sexes. These distributions are presented in Appendix A. While the largest number of Ss of both sexes fell within the CAs of 15-19 years, the percentage of Ss in this age range was even higher for males than for females. This male population was in general younger than this female population, with over two-thirds of the males 19 years of age or younger, while fewer than half of the females were 19 years of age or younger. Conversely, 22.9% of the female Ss were in the 30-34 age range, while only 2.9% of the male Ss fell into this range. The fact that the male Ss were as a whole younger than the female Ss was clearly shown by the mean CA figures for the 68 male and 48 female eligible Ss; the mean CA for males was 19-1.5, while the mean CA for females was 23-2.3.

The IQs of the males and females in the pool of eligible Ss were more evenly distributed and were equivalent for both sexes. The mean IQ for the 68 male Ss was
60.2 and for the 48 female Ss 60.3. The distribution of IQs for the entire eligible subject pool is presented in Appendix A.

All the IQ data were based on testing done at Denton State School during the past five years. All but four of the IQ scores represented the Full Scale IQ on the Wechsler Adult Intelligence Scale (WAIS) or the Wechsler Intelligence Scale for Children (WISC); the exceptions were scores on the Stanford-Binet Intelligence Scale, Form L-M.

Subject Selection

Selection of the 35 male and 35 female Ss for the experimental groups was done by means of the following procedure. Index cards on each S were separated according to sex and were arranged according to ascending CA. The cards were then divided into seven groups, and experimental groups were composed by means of random selection of one S from each of the seven groups of cards. This procedure insured an even CA distribution in each experimental group. However, due to the differences in CA distribution between the two sexes, it was necessary to systematically substitute certain Ss in all groups in order to increase the mean CA of the male groups and decrease the mean CA of the female groups.

The mean CAs for the experimental groups varied for male Ss from 22-7.14 to 22-8.57 and for female Ss from
22-7.57 to 22-9.00. The mean IQ for all experimental groups, both male and female, ranged from 60.43 to 61.43. Thus, the experimental groups were well matched; there were no significant differences in mean CA or mean IQ between any of the experimental groups or between the sexes in these groups. The mean CAs and IQs for all experimental groups can be found in Appendix A.

Assignment of groups of Ss to experimental conditions was made on an arbitrary basis.

**Apparatus**

The apparatus was designed to perform the same functions as the apparatus employed in the Bandura and Perloff (1967) study. All parts of the apparatus to be manipulated by the Ss were mounted into the front panel, a rectangular piece of masonite 19 inches high and 21-3/4 inches wide. Masonite panels were also placed on the top and either side of the apparatus to protect the interior components and to hide them from the Ss' view. The back of the apparatus was open to provide the access to the dispenser and remainder of the components.

Mounted into the front panel were the following items:

1. A five-inch crank which was manually operated by the Ss.
2. A criterion-selector switch which could be set on Off, or on 5, 10, 15, or 20 points.

3. A vertical panel divided into four translucent rectangular sections (3 inches by 3-1/2 inches) upon which were painted the numerals 5, 10, 15, and 20 (indicating performance levels) in ascending order. These sections lighted up successively as the S achieved the corresponding criterion levels by turning the crank the required number of times and extinguished simultaneously when the final criterion was achieved.

4. A small chute through which tokens were dispensed into a token receptacle, a transparent plastic cup which was placed on the table directly in front of and below the token dispenser chute. Both the chute and the token receptacle were enclosed in a transparent plastic box which allowed the Ss to view the tokens but prevented them from handling them or taking them out of the token receptacle.

5. A self-reinforcement button located directly above the token dispenser chute. Ss in the self-reinforcement conditions dispensed tokens to themselves when they had achieved their performance criterion by pushing this button.
The dispenser used was an M & M dispenser. Since no other type of token could be found which did not jam in the dispenser, M & Ms were sprayed with a beige enamel paint so that they lost their appearance of being candy and functioned only as tokens.

The apparatus also had a buzzer which sounded when the performance criterion was reached. The ratio of crank revolutions : points achieved was set so that eight revolutions of the crank were required to achieve a 5-point score and eight more revolutions to achieve each subsequent score. Thus, 8 turns of the crank were necessary to achieve a score of 5, 16 turns for a score of 10, 24 turns for a score of 15, and 32 turns for a score of 20. When a high criterion was set, panels representing lower performance scores lighted up in ascending order as the S approached the criterion level which had been set. For example, if the criterion was set at 20, the 5-, 10-, and 15-point score indicators lighted up when the crank had been turned 8, 16, and 24 times, respectively. When the S turned the crank a total of 32 times, the 20-point score indicator lighted up, a buzzer sounded, and all score lights extinguished simultaneously. Then, depending on the experimental condition, tokens were automatically dispensed to the S or he pushed the button to dispense the tokens to himself.
Connected to the apparatus described above by means of a cable was a remote control panel which was located in an adjacent observation room and was operated by E. On this panel were located the following items:

1. The counter, which automatically recorded the number of revolutions of the crank, and a counter reset mechanism.

2. A switch which controlled the source of reinforcement, self-reinforcement or automatic (external) reinforcement.

3. A criterion control switch, whose setting determined whether the performance criterion was set by S or by E.

4. A criterion-selector switch for use in conditions in which the criterion was determined by E.

5. Score-indicator lights for performance scores of 0, 5, 10, 15, and 20.

6. A light which indicated when the S had achieved his criterion.

7. A reward indicator which lighted each time a token was dispensed.

8. A reward button by means of which E could manually dispense tokens at any time.

9. A reset button by which E returned the apparatus to zero (the beginning of a new run).
10. A step button by means of which E could advance the mechanism one step at a time, just as the S advanced it by turning the crank.

Diagrams of the front panel of the apparatus and the experimenter's remote panel are presented in Appendix B. Power was supplied by a Nobatron DCR 40-10A Power Supply, adjusted to a 28-volt setting.

**Pilot Study A: Rankings of Reinforcement Values**

This pilot study was conducted in order to obtain rankings of the reinforcement values of the various reinforcers to be used with the experimental groups.

**Procedure**

The Ss were five male and five female institutionalized mental retardates whose CAs and IQs corresponded as closely as possible to those of the experimental Ss. The mean CA of the male Ss was 21-8.40, and their mean IQ was 61.00. The female Ss had a mean CA of 23-4.00 and a mean IQ of 60.60.

The reinforcers ranked in the pilot study consisted of 26 consumable and nonconsumable items of low monetary value (ranging from 5¢ to 29¢). The reinforcers chosen to be ranked in the pilot study were the following: gum, candy bar, Fritos, Cracker Jacks, M & Ms, peanuts, gum drops, cupcakes, toy watch, comb, plastic truck, water

Each S was told that the experimenter wanted to find out what kinds of things he or she liked and was then asked, "If you could have one of these things, which one would you most want to have?" When the S had made a choice, it was recorded, that item was removed, and the procedure was repeated until all items had been selected and a complete ranking of reinforcement values had been obtained. This procedure was modeled after that used by Heber (1959). Following this procedure each S was asked, "Can you think of any other things you would like to have that were not on the table?" and responses were recorded. The S was then given the item which he had ranked as his first choice and was thanked for his cooperation.

Results

The mean rank of reinforcement value was computed for each of the 26 reinforcers. There were great interindividual differences in preferences, as indicated by the facts that no item received a mean ranking smaller than 8.6 or larger than 18.9 and that there were eight different items chosen as most preferred by the ten Ss. Reinforcers were listed according to their mean rank values, and the
item having the median reinforcement value was determined by eliminating the item of lowest mean rank so that an odd number of items (25) remained and the median rank was the thirteenth item on the list. This item, which was to serve as the experimenter-selected reinforcer in three experimental conditions, was the bag of Fritos, which had a mean rank reinforcement value of 12.7. The mean rankings of the items for the five male and five female Ss combined appear in Appendix C.

The original intention of the experimenter was to eliminate any item having a very low mean ranking in the pilot study from the reinforcer pool to be used in the experimental conditions featuring subject selection of reinforcer. However, since no items had extremely low mean rankings none were eliminated from the reinforcer pool. It was also intended that additional items reported by Ss as being desirable reinforcers would be added to the reinforcer pool for the experimental conditions. All Ss except one reported additional desirable reinforcers and many suggested numerous items; however, a majority of the items mentioned were too expensive to use, and food items mentioned were either very similar to items already in the reinforcer pool (as a different kind of candy bar or chip) or were items which required refrigeration (soft drinks and ice cream). Although a few items were mentioned which could have been added to the reinforcer pool, including puzzles,
nail polish and nail polish remover, sugar; newspapers, magazines, and TV Guide, since these items were so few in number and since all Ss in the pilot study indicated definite liking for at least some of the items in the original reinforcer pool, it was decided not to incorporate these items into the reinforcer pool. Retaining the same reinforcer pool that was used in the pilot study assured that the experimenter-selected reinforcer used in the experimental conditions would be the one having the median rank of reinforcement value for that particular pool of reinforcers.

**Pilot Study B: Change Option in Setting of Performance Criterion**

The purpose of this pilot study was to determine whether or not it would be feasible with the current population of mentally retarded Ss to introduce the change option employed by Bandura and Perloff (1967), whereby Ss in the self-reinforcement conditions were allowed to change their performance criterion once during the experiment. If non-experimental mental retardates were able to understand and follow these additional instructions in the pilot study, the change option was to be used in the experiment; if not, the change option was to be omitted and Ss in the self-reinforcement conditions would be allowed to make only one choice of performance criterion.
This pilot study was designed to include a total of six Ss, three males and three females, whose CAs and IQs corresponded as closely as possible to those of the experimental Ss.

Procedure

Ss were familiarized with the apparatus by means of verbal description, a demonstration trial by E at one criterion level, and a practice trial by S at a different criterion level. The Ss were told that when E left the room they were to select the number of points for which they wished to work by setting the criterion selector switch at that number and that if they wished to do so they could change the criterion one time only during the experiment. They were also instructed that when they achieved the criterion and the buzzer sounded they could take as many tokens as they liked, and that if they earned enough tokens they would get a "prize," a bag of Fritos. They were told they could work as long as they liked at the machine and could stop whenever they wanted.

Results

It was necessary to run only two of the six pilot study Ss in order to determine conclusively that the change option could not be used with this population of institutionalized mentally retarded Ss. Two male Ss were run,
both of whom blatantly violated the instructions to reset the criterion only one time. The first S set his initial criterion at 5 points, then switched it to 10 points. However, with the criterion set at 10, when 5 points were earned and tokens were not available, the S switched the criterion back to 5 and repeated this same procedure several times. The second S adjusted his criterion setting in an even more capricious manner, changing his criterion innumerable times and frequently within a run, before any criterion level had been achieved. These multiple changes of criteria on the part of both of the pilot study Ss made it apparent that the change option employed by Bandura and Perloff (1967) would have to be eliminated from the current study.

As a result of the findings of this pilot study, Ss in the self-reinforcement conditions were instructed that they were to choose their performance criterion and that once the criterion was selected they were not to change it. Eliminating the change option made it possible to switch control of the criterion from the S's control to the E's control once the criterion had been set by the S. In this manner the criterion initially selected by the S remained constant throughout the experimental procedure even though the S might change the criterion switch on the apparatus to a different setting.
**Experimental Procedure**

All Ss were run in audiological testing rooms which were located in the Academic School Building at Denton State School. Use of these particular facilities offered several advantages: The cable connecting the main part of the apparatus to the remote control panel could be run through the wall between the two rooms along with the wiring of the audiological testing equipment, the E could observe the S through a one-way observation window, and the rooms were soundproof and without any windows so that outside distractions to the Ss were eliminated.

Each S was familiarized with the apparatus and given instructions appropriate to his particular experimental group.

**Group A: Self-reinforcement, subject-selected reinforcer group.** Each S was brought into the room and seated directly in front of the apparatus. The E explained and demonstrated how the apparatus worked by going through all the steps at one performance criterion level. Then the S was given a practice trial with the apparatus set at a different criterion. It was emphasized that when E left the room the S could set the apparatus at whatever criterion he chose but that he was not to change the criterion he had initially selected. Each S was told that whenever he achieved his criterion he could take as many tokens as he
liked and that if he earned enough tokens he could have his choice of any of a collection of 26 backup reinforcers which were displayed on a table in the same room. Each S was told that he could work as long as he liked and stop whenever he wished.

**Group B: Self-reinforcement, experimenter-selected reinforcer group.** The instructions were the same for this group as for Group A, with the exception that Ss in this group were shown only the experimenter-selected reinforcer (a bag of Fritos) and were told that if they earned enough tokens they would get it as their prize.

**Group C: External reinforcement, subject-selected reinforcer group.** The performance criterion levels and the magnitude of token reinforcement for the Ss in this group were yoked to those selected by the Ss in Group A. The yoking for each S was with an S in Group A of the same sex and in the same position within the experimental group according to ascending CA. Since the criterion was predetermined, all instructions and demonstrations were carried out with the apparatus set at that criterion. As in the other groups, the apparatus was demonstrated by E, followed by a practice trial by the S. Ss were told that the number of tokens they received each time they reached the criterion would be determined by the apparatus. They were shown the collection of 26 backup reinforcers and told that if they earned enough tokens, they could have their choice of these
items. As in the other groups, Ss were told they could work as long as they liked and stop whenever they wished.

**Group D: External reinforcement, experimenter-selected reinforcer group.** The instructions and demonstration for Ss in this group were the same as for the Ss in Group C, with the exception that these Ss were shown only the experimenter-selected reinforcer (a bag of Fritos) and told they would win it if they earned enough tokens. Performance criteria and magnitude of token reinforcement were yoked with Group B.

**Group E: Noncontingent reinforcement control group.** Ss in this group were yoked to Ss in Group B for performance criterion and total number of tokens received. However, Ss in this group were given the total number of tokens earned by the Ss to whom they were yoked prior to beginning the task and were told that the tokens could be exchanged for the experimenter-selected reinforcer (a bag of Fritos) when they finished performing. Ss in this group worked for a criterion, and the score indicators lighted up and the buzzer sounded when the criterion was reached, but no tokens were dispensed contingent upon this performance. Therefore, the Ss were instructed that they were to work for a certain number of points throughout the procedure, and the apparatus was set at this criterion for both the E's demonstration and the S's practice trial. As in the other
experimental groups, Ss were told that they could work as long as they wanted and stop whenever they liked.

Ss in all conditions performed in the absence of E. All Ss were observed by E through a one-way observation window, and performance data were recorded on sheets designed for this purpose. Copies of these Performance Recording Sheets for self-reinforcement, external reinforcement, and noncontingent reinforcement control conditions are found in Appendix D. The major dependent variable was maintenance of behavior, as measured by the total number of revolutions of the crank for each S. In the self-reinforcement conditions, criterion set and number of tokens taken were recorded for purposes of yoking and comparison with the data yielded in Bandura and Perloff's (1967) study with normal children.

After each S had indicated by leaving the testing room that he had finished performing, he and E returned to the testing room where the S either selected the reinforcer of his choice or was given the experimenter-selected reinforcer. All Ss received a backup reinforcer regardless of the actual number of tokens earned during the experimental procedure.

Following receipt of their backup reinforcers all Ss were asked a series of questions designed to determine
their concepts and expectations regarding the experimental task. Copies of the Experimental Demand Questions for the self-reinforcement, external reinforcement, and noncontingent reinforcement control groups are found in Appendix D.
CHAPTER 4

RESULTS

The mean performance scores (number of crank revolutions) of the experimental and control groups are presented in Table 1.

Table 1. Mean Number of Crank Revolutions in Experimental and Control Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Number of Crank Revolutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Self-reinforcement, subject-selected reinforcer group</td>
<td>577.86</td>
</tr>
<tr>
<td>B: Self-reinforcement, experimenter-selected reinforcer group</td>
<td>767.64</td>
</tr>
<tr>
<td>C: External reinforcement, subject-selected reinforcer group</td>
<td>582.00</td>
</tr>
<tr>
<td>D: External reinforcement, experimenter-selected reinforcer group</td>
<td>798.93</td>
</tr>
<tr>
<td>E: Noncontingent reinforcement control group</td>
<td>642.21</td>
</tr>
</tbody>
</table>
A 2 X 2 X 2 analysis of variance (Edwards, 1960) was computed on the performance scores (number of crank revolutions) of Ss in the four experimental groups to determine the significance of main effects and interactions involving the following factors: (a) type of reinforcement, self-reinforcement versus external reinforcement, (b) source of reinforcer selection, subject selection versus experimenter selection, and (c) sex, male versus female. It was hypothesized that Ss in the subject-selected reinforcer groups would have significantly higher performance scores than Ss in the experimenter-selected reinforcer groups, that Ss in the external reinforcement conditions would have higher performance scores than Ss in the self-reinforcement conditions, and that males would have higher performance scores than females. None of the F values yielded in this analysis were significant at the .05 level, indicating that number of responses was not a differential function of type of reinforcement, source of reinforcer selection, sex, or any interactions between these factors. Thus, none of the above hypotheses were supported. A summary of this analysis is presented in Table 2.

Differences in the mean scores of the four experimental groups were compared, with sexes both combined and separated, by means of Duncan's New Multiple Range Test (Edwards, 1960). The results indicated that there were no significant differences between any of the group means.
Table 2. Complete 2 X 2 X 2 Analysis of Variance of Performance Scores (Number of Crank Revolutions) of Experimental Groups

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Type Reinforce</td>
<td>104,393.12</td>
<td>1</td>
<td>104,393.12</td>
</tr>
<tr>
<td>B</td>
<td>Source Reinforcer Selection</td>
<td>578,957.75</td>
<td>1</td>
<td>578,957.75</td>
</tr>
<tr>
<td>C</td>
<td>Sex</td>
<td>132,697.75</td>
<td>1</td>
<td>132,697.75</td>
</tr>
<tr>
<td>A X B</td>
<td>Type X Source</td>
<td>2,578.57</td>
<td>1</td>
<td>2,578.57</td>
</tr>
<tr>
<td>A X C</td>
<td>Type X Sex</td>
<td>244,464.29</td>
<td>1</td>
<td>244,464.29</td>
</tr>
<tr>
<td>B X C</td>
<td>Source X Sex</td>
<td>331,716.07</td>
<td>1</td>
<td>331,716.07</td>
</tr>
<tr>
<td>A X B X C</td>
<td>Type X Source X Sex</td>
<td>81,033.15</td>
<td>1</td>
<td>81,033.15</td>
</tr>
<tr>
<td>Error</td>
<td>Within Treatments</td>
<td>20,671,236.60</td>
<td>48</td>
<td>430,650.76</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>22,147,077.30</td>
<td>55</td>
<td></td>
</tr>
</tbody>
</table>

* >.05
Analysis of variance and Duncan's New Multiple Range Test analyses of square root transformations of the original performance scores yielded results which were not appreciably different from the same analyses of the original performance scores.

It was hypothesized that the noncontingent reinforcement control group would have the lowest performance of all groups. Dunnett's Test for Comparisons with a Control (Edwards, 1960) indicated that there were no significant differences between the control group mean and the means of any of the experimental groups. However, comparison of the control group with the experimental groups was complicated by a significant difference between males and females in the control group ($t = 2.57, p < .025$ with original performance scores, $t = 3.06, p < .005$ with square root transformations). Therefore, further analyses were conducted, comparing control group males and control group females separately with the experimental groups. The results indicated that control group females had significantly lower performance scores than Group B ($t = 3.05, p < .005$), Group C ($t = 1.93, p < .05$), and Group D ($t = 2.79, p < .01$). The difference between control group females and Group A was in the expected direction but achieved significance only at the .10 level ($t = 1.38$). The $t$ tests comparing control group females with experimental groups using square root transformations of scores
indicated that performance of control group females was significantly lower than Group A ($t = 2.12, p < .025$), Group B ($t = 3.83, p < .005$), Group C ($t = 2.53, p < .025$), and Group D ($t = 3.51, p < .005$). The results for control group males were quite different; the mean for control group males exceeded the means of all the experimental groups. While none of the differences between the means of control group males and experimental groups were significant at the .05 level, the differences between control group males and Groups A and C approached significance ($t = 1.38, p < .10$ and $t = 1.62, p < .10$, respectively). Analyses with square root transformations of scores yielded the same results. The $t$-test analyses on the original performance scores are presented in Table 3. Thus, the hypothesis that the control group would have significantly lower performance than the experimental groups was supported for the females in the control group but not supported for the males in the control group.

Bandura and Perloff (1967) found that external reinforcement resulted in greater performance than self-reinforcement with boys only. In order to determine whether this was true in the current study, a comparison was made between the performance of males in the two self-reinforcement groups and the two external reinforcement groups. However, the difference in performance of these
Table 3. Comparisons of Control Group Means (Male and Female) with Means of Experimental Groups

<table>
<thead>
<tr>
<th>Groups Compared</th>
<th>t</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>A and Control Females</td>
<td>1.38</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>B and Control Females</td>
<td>3.05</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>C and Control Females</td>
<td>1.93</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>D and Control Females</td>
<td>2.79</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Control Males and A</td>
<td>1.38</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>Control Males and B</td>
<td>1.12</td>
<td>&lt;.15</td>
</tr>
<tr>
<td>Control Males and C</td>
<td>1.62</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>Control Males and D</td>
<td>.96</td>
<td>&lt;.20</td>
</tr>
</tbody>
</table>
two groups was not significant ($t = .70$). Nor was there a significant difference for females between self-reinforcement and external reinforcement ($t = .43$).

The primary factor contributing to the lack of significance found in the analysis of variance of the performance scores of the experimental groups appeared to be the large variance within groups. An examination of the distribution of performance scores revealed that while 49 (87.5%) of the scores fell within the 1-1500 range, 7 (12.5%) were scattered above this range with the highest a score of 3333. The Performance Recording Sheets of Ss with scores over 1500 were examined in an attempt to identify factors contributing to such high scores. The entire range of CAs and IQs were represented, and high-scoring Ss appeared in all experimental groups. However, of the 7 Ss making over 1500 responses, 5 had contact with $E$ subsequent to the initial instructions. Such contact was of two types: (a) $S$ seeking out $E$ during the experimental procedure to ask a question or make a complaint about the apparatus; such questions and complaints were attributable to Ss' failure to understand some part of the instructions given, and (b) $E$ interrupting the experimental procedure to unjam the dispenser. These Ss were not eliminated from the experimental groups due to the unavailability of suitable replacements and $E$'s faulty judgment that such deviations
in the standard experimental procedure would not significantly affect performance. Performance Recording Sheets of all Ss were examined for similar occurrences of Ss having contact with E subsequent to the initial instruction period. Seven additional instances were found among Ss who were not in the high-scoring group, for a total of twelve Ss who had extra E-contact. Of these Ss, 4 were in Group A, 6 in Group B, 1 in Group C, and 1 in Group D; the high proportion in Groups A and B was attributable to the fact that all instances of Ss seeking out E to ask a question or make a complaint occurred in these self-reinforcement groups. There were no differences in mean IQ or CA between these 12 Ss and the remaining 44 Ss; Ss having extra E-contact had a mean CA of 22-10.25 and a mean IQ of 61.75, while Ss not having extra E-contact had a mean CA of 22-7.64 and a mean IQ of 60.64. A t test of the mean scores of these two groups (1239.58 and 529.43) yielded a highly significant difference (t = 3.84, p < .005 with original performance scores, t = 3.69, p < .005 with square root transformations), indicating that the extra E-contact during the experimental procedure significantly increased performance.

Because of the significantly higher performance scores of Ss having extra E-contact, attempts were made to analyze the data on the remaining 44 Ss who had experienced
no extra E-contact. Analysis of variance was ruled out due to unequal ns of the experimental groups after elimination of the 12 extra E-contact Ss, and t tests were computed comparing group means. The only significant difference found was between Groups A and D ($t = 1.98, p < .05$ with original performance scores, $t = 1.85, p < .05$ with square root transformations). This comparison was confounded by the fact that groups were not matched for performance criteria and magnitude of reinforcement. In an attempt to rectify this situation, Ss who were yoked to extra E-contact Ss were also eliminated, leaving a total of 32 Ss in the experimental groups. Comparison of the means of these revised experimental groups by means of t tests yielded no differences which were significant at the .05 level. These data are, at best, inconclusive.

The final hypothesis made was that Ss in the self-reinforcement conditions of the current study would select less stringent performance criteria and would self-administer a greater magnitude of token reinforcement/reinforcement period than did the Ss in the self-reinforcement group in Bandura and Perloff's (1967) study. In the current study the demonstration and practice trials involved criteria of 5 and 15, respectively. However, there did not appear to be an imitation effect resulting in Ss' frequent selection of these settings. Of the 28 Ss in the self-reinforcement groups, 15 (53.57%) selected a criterion of
20, 3 (10.71%) selected a criterion of 15, 5 (17.86%) selected a criterion of 10, and 5 (17.86%) selected a criterion of 5. Thus, a criterion of 20 was selected by slightly over half of the $S$s, with the remaining $S$s selecting a criterion of 5, 10, or 15. Of the 20 $S$s in Bandura and Perloff's self-reinforcement group, 9 set a criterion of 20, 7 set a criterion of 15, and 4 a criterion of 10, but none of the $S$s selected the lowest performance criterion of 5. The performance criteria selected by $S$s in the current study and in the Bandura and Perloff study both varied significantly from change distribution (in which one-fourth of the $S$s would have selected each criterion).

Chi-square comparison of expected and obtained frequencies of criteria selected yielded a $\chi^2 = 12.57 (p < .01)$ for the current study and a $\chi^2 = 9.20 (p < .05)$ for the Bandura and Perloff study. Comparison of these two sets of data does not lead to conclusive statements about the differences between the two populations. Roughly half of the $S$s in both studies selected the highest performance criterion. If the two highest and two lowest criteria are combined, the hypothesis that normal $S$s select more stringent criteria appears to gain some support; in the Bandura and Perloff study 80% of $S$s chose one of the two higher criteria and only 20% chose one of the lower criteria, while in the current study 64% of $S$s chose one of the higher criteria and
36% chose one of the lower criteria. However, when the data from the two studies were submitted to chi-square analysis, the difference between them failed to attain significance at the .05 level ($\chi^2 = 1.39$).

An important factor which must be considered in comparing the performance criteria set in the two studies is the matter of changing criteria. Bandura and Perloff employed a change option whereby Ss could make one change in criterion set during the experimental procedure; 16 of 20 Ss did change the criterion set, 6 to a higher criterion and 10 to a lower criterion. In the current study, however, it was necessary to eliminate the change option, allow Ss only one choice of criterion, and thereafter switch control of criterion to the E's remote panel. In spite of very explicit instructions not to change the criterion once it was set and the fact that altering the S's criterion switch did not result in corresponding change in criterion to be achieved before tokens were available, 19 of 28 Ss in the two self-reinforcement groups moved the criterion switch to a different setting at least once after original selection of criterion had been made. Of these 19 Ss, 5 changed the criterion only one or two times and returned it to the original setting, apparently having noted that changes in the criterion setting subsequent to the original setting did not change the functioning of the
apparatus. The remaining 14 Ss (half of the total number of Ss in the self-reinforcement groups) changed the setting of the criterion switch numerous times during the experimental procedure, generally at the beginning of a run but sometimes in the middle of a run. The presence of such behavior tends to indicate that setting the criterion switch at a particular level did not have the meaning of setting an achievement standard for these Ss. Therefore, any potential meaning that the above data on criterion selection might have is greatly diminished. As Bandura and Perloff made no mention of such a phenomenon, presumably it did not occur in their Ss.

The hypothesis that Ss in the current study would self-administer a higher magnitude of token reinforcement than did Ss in the Bandura and Perloff study was confirmed. Bandura and Perloff's Ss took an average of 1.02 tokens/reinforcement period, regardless of the criteria for which they were working. The overall mean number of tokens taken/reinforcement period for the 28 Ss in the self-reinforcement conditions in the current study was 8.18, with mean number of tokens self-administered/reinforcement period for individual Ss ranging from .90 to 30.88. The difference between the means of the current study and Bandura and Perloff's study was statistically significant ($t = 1.98, p < .05$).
There did not appear to be any significant relationship between criterion selected and magnitude of self-administered token reinforcement in the current study; for Ss working for a criterion of 20 the mean number of tokens taken/reinforcement period was 8.11, for a criterion of 15 the mean was 7.30, for a criterion of 10 the mean was 7.14, and for a criterion of 5 the mean was 9.97. In addition, analysis of variance indicated that in the self-reinforcement groups in the current study there were no significant differences in average magnitude of token reinforcement as a function of sex, source of reinforcer selection, or an interaction between these two factors. Therefore, it can be concluded that institutionalized mentally retarded are more generous than normal children in self-administering reinforcement, but that such generosity is not related to performance criterion set, sex, or source of reinforcer selection.

Additional analyses were made of the performance of experimental and control groups. Due to the wide range in CA of the Ss (from 12-8 to 34-6) an attempt was made to determine whether CA and performance were significantly related. All Ss in the experimental and control groups were divided into two groups according to age, Ss age 20 and below and Ss age 21 and above. Comparisons of the original performance scores and of the square root transformations
of these scores by means of $t$ tests indicated that there was no significant difference between these two age groups ($t = .46$ and $t = .74$, respectively).

Since the various performance criterion levels actually represented different fixed ratio schedules of reinforcement--FR8, FR16, FR24, and FR32--the performance data were analyzed to determine whether there were significant differences in total number of responses for Ss working for the four different performance criteria. $t$ tests indicated that there were no significant performance differences between any two criteria in the self-reinforcement groups, external reinforcement groups; self-reinforcement and external reinforcement groups combined, or in all experimental groups and the control group combined.

Analysis of variance of the total number of runs completed to criterion yielded results which were comparable to those of the analysis of variance of number of crank revolutions. There were no main effects or interactions which were significant at the .05 level.

Total number of tokens obtained by the experimental groups (either by self-reinforcement or external reinforcement) was submitted to a $2 \times 2 \times 2$ analysis of variance. There were no significant main effects or interactions.
Ss' responses to the Experimental Demand Questions indicated that they had little, if any, idea of the purpose of the experiment. All Ss were asked, "What do you think this study is about?" Twenty-four Ss (34%) responded only that they did not know, 14 (20%) thought it was some type of test of their intelligence, 11 (16%) thought it had something to do with learning, and 7 (10%) considered it to be some sort of game. Only 10 Ss (14%) emphasized earning tokens or points to get a prize, and many of these responses were vague ("Tokens--to buy stuff with," and "Win some points, I guess"). The best responses in this category included "It's about tokens, how you get the most points" and "Tokens work like merits, you spend them the same way."

When asked what they thought they were supposed to do, 28 of the Ss (40%) simply repeated part of the instructions that had been given to them, and 19 (27%) indicated that they did not know. Of a total of 13 Ss (19%) whose responses involved the idea of earning tokens or points, only 2 Ss (3%) mentioned a relationship between quantity of tokens and getting a prize ("See how many tokens to win a prize" and "Get all that up to the top and get the Fritos").

In addition to the above questions, all Ss were asked why they stopped turning the crank when they did and
how long they thought they were supposed to turn it. Twenty-five Ss (36%) said they stopped because they were tired, 6 (9%) said they did not know why they stopped, and 10 (14%) gave no response or an unintelligible, inappropriate, or nonspecific response (such as "I just wanted to"). Seven Ss (10%) stopped because they had obtained all available tokens and 13 (19%) stopped because they thought they had enough tokens to win the prize. Of the latter group, 5 Ss (7%) made the decision on the basis of how full the token receptacle was. In spite of this total of 20 Ss (29%) who indicated that they stopped because they had all or enough tokens, only 7 Ss (10%) indicated that they thought they were supposed to keep turning the crank until they reached a certain goal, such as getting enough tokens. Instead, 31 Ss (44%) indicated that they did not know how long they were supposed to keep turning the crank, 11 (16%) gave a response involving a specific or vague amount of time, 4 (6%) said until they got tired, 4 (6%) said as long as they wanted to, and 8 (11%) did not answer or gave an unintelligible or inappropriate response.

Ss in the experimental groups were asked how many tokens they thought they needed to get the prize. Fourteen Ss (25%) indicated they did not know, and 3 (5%) indicated an unspecific quantity (such as "a lot"). Twenty-six Ss (46%) reported a specific number, ranging from 4 to 200;
the striking feature of these responses was their random nature and the Ss' inability to state why they thought they needed that particular number of tokens. Eleven Ss (20%) indicated that they thought they needed all of the tokens, the amount they had earned, or some amount relative to the size of the token receptacle. In spite of this variability in responses regarding how many tokens Ss thought they needed to win the prize, 47 Ss (84%) gave an affirmative response to the question, "Did you think you had enough tokens to get the prize?" These responses may have been influenced by the fact that Ss had received their backup reinforcers before they were asked the Experimental Demand Questions.

Ss in the self-reinforcement groups were asked additional questions about the performance criteria and magnitudes of token reinforcement that they had selected. In response to the question of why they had selected the performance criterion they did, 4 Ss (14%) indicated they did not know why, 3 (11%) said they "just wanted to," 8 (28%) stated that they liked that number or that it was a lucky number, and 4 (14%) gave no response or an inappropriate one. Seven Ss (25%), all of whom selected a criterion of 20, indicated that they selected that criterion because it was the highest. Some of these responses implied that because 20 was the highest criterion, it yielded more
tokens ("It's the highest; you have to put it on the highest to win" and "You get more tokens on 20 than on 15"). When asked where they thought they were supposed to set the criterion switch, 4 Ss (14%) said on any number they wanted, 4 (14%) said they did not know, and 2 (7%) mentioned several criteria in a sequence. The majority of Ss--18 or 64%--reported one of the criteria, but only 4 of these Ss (14%) could give any reason for their responses. As with the responses to the question of how many tokens were needed to win a prize, the responses to this question seemed to be random rather than reflections of Ss' expectations of the experiment.

Responses to the question of why Ss took a certain number of tokens yielded even less information as to the determinants of the Ss' behavior. Ten Ss (36%) did not know why they took the number of tokens they did, 5 (18%) "just did" or "just wanted to" take a certain number, and 9 (32%) gave unintelligible or meaningless responses. Four Ss (14%) reported taking a certain number of tokens each time in order to get enough tokens to win a prize. None of the Ss related magnitude of token reinforcement to their performance criterion in any way. The question "How many tokens did you think you were supposed to take each time?" resulted in more random responses. While 6 Ss (21%) gave no response or indicated they did not know and 1 (4%) said
he was supposed to take as many tokens as he wanted to, 21 (75%) of the Ss gave a specific number, ranging from 1 to 20. Some of these Ss gave more than one response and others gave estimates. None of the Ss were able to state a reason for their responses.
DISCUSSION

Before the results of the current study and their implications are discussed, it must be emphasized that any conclusions apply only to institutionalized retardates. Several writers (Gerjuoy and Winters, 1969; Kaufman, 1967) have warned against generalizing the results of studies on institutionalized retardates to noninstitutionalized retardates. Hobbs (1964) compared groups of institutionalized and noninstitutionalized retardates of equivalent CA and IQ and found many differences between the two groups, which were viewed as variables contributing to institutionalization. Data on institutionalized retardates may not be representative of all retardates because of the selective factors operative in institutionalization and the unique environment in which institutionalized retardates live.

While Ss in the current study were not representative of all retardates due to the fact of their institutionalization, these Ss were very likely representative of mild mental retardates who are institutionalized. The sex and age distributions of the eligible subject pool (the latter presented in Appendix A) were in agreement with the findings of Churchill (1964) that among admissions of mild
mental retardates to institutions, males were admitted more frequently and at a younger age than females.

The results of the current study indicated that there were no significant main effects of type of reinforcement (self-reinforcement versus external reinforcement), source of reinforcer selection (subject versus experimenter), or sex; nor were there any significant interactions among these factors. Various reasons for the lack of difference between the two sources of reinforcer selection can be hypothesized. It can be argued that due to the paucity of reinforcers available to institutionalized retardates, they fail to discriminate between the reinforcement values of various items, hence any reinforcer is effective in maintaining behavior in this population. The fact that performance in all experimental groups was higher than the performance of females in the control group (who received only noncontingent reinforcement) supports this hypothesis, but the high performance scores of males in the control group indicate that this may not have been the only factor involved. Another possible reason that no differences were found between conditions of subject selection and experimenter selection of reinforcers could be the low value of the reinforcers used. Reinforcers were restricted in type to material reinforcers and in monetary value to 29¢ or less. Even if institutionalized mental retardates do
discriminate between the reinforcement values of various items, there may not have been a great enough differential in reinforcement values among the reinforcers used in the subject selection conditions to insure their having a greater effect on behavior than an experimenter-selected reinforcer. The fact that Ss in the subject-selected reinforcer groups chose items other than the bag of Fritos (the experimenter-selected reinforcer) as their backup reinforcers indicated that these reinforcers had greater value; however, it was impossible to evaluate the magnitude of this differential value in the current study.

There were several sources of reinforcement common to all groups, including the control group, which may have served to equalize the differential effects of source of reinforcer selection and type of reinforcement. All of the following factors may have had such reinforcing effects: The fact of being selected as an S and the opportunity to engage in an activity which differed from the everyday schedule of the Ss, the social contact with E during the instruction period, and the lighting up of the score panel and the buzzer sounding to indicate that a criterion had been reached. In addition, all experimental groups received token reinforcement. Since the current study was not designed to evaluate the differential reinforcing
effects of each of these factors, only speculations can be made about their effects.

Gerjuoy and Winters made the statement that "the retardate's participation in an experiment may stand out as an important event in his day" (1969, p. 9). Such a statement implies that such participation has reinforcing properties for the institutionalized retardate. Certainly the finding that Ss having additional contact with E subsequent to the instruction period had significantly higher performance scores than did Ss having no extra E-contact emphasized the important influence of such contact. While attempts were made to eliminate social reinforcement from the current study, including E observing S from an observation room rather than being present during the experimental procedure, these attempts were not successful in the cases of the 12 Ss who had extra E-contact. The fact that this contact was a major confounding factor, resulting in significantly higher performance scores for Ss experiencing extra E-contact, might have been predicted from the literature on the effects of social reinforcement on institutionalized retardates. Because of its importance and relevance to the results of the current study, this literature will be reviewed briefly.

Several studies have established social reinforcement as being very effective in maintaining behavior in
institutionalized retardates. Zigler based his hypotheses concerning the effects of social reinforcement on institutionalized retardates on the assumption that "institutionalized feeble-minded children tend to have been relatively deprived of adult contact and approval and hence have a higher motivation to secure such contact and approval than do normal children" (1961, p. 413). Results of several studies have supported Zigler's hypotheses, indicating that institutionalized retardates will spend more time at a monotonous task under conditions of social reinforcement than will normals or noninstitutionalized retardates (Green and Zigler, 1962; Zigler, Hodgden, and Stevenson, 1958). The fact that there were no significant differences between normals and noninstitutionalized retardates on this variable in Green and Zigler's study led them to conclude that "the greater perseverance of the institutionalized retarded child on satiation type tasks can most parsimoniously be attributed to the particular environmental conditions he has experienced rather than to his intellectual level" (1962, p. 505). The results of Stevenson and Fahel's (1961) study supported this viewpoint; they found social reinforcement to be more effective in maintaining the behavior of institutionalized normal and retardates than of noninstitutionalized normals and retardates. Shepps and Zigler (1962) found no differences between familial and
organic retardates in the effect of social reinforcement; both types of retardates responded longer under conditions of social reinforcement than under no reinforcement.

Zigler, et al. (1958) did note, however, that there were "large individual differences among feebleminded Ss in their susceptibility to the effects of support" (p. 121). Zigler (1961) hypothesized that such variation was a function of the degree of preinstitutional deprivation and found that Ss rated as having a highly deprived preinstitutional history spent significantly longer on a monotonous task under conditions of support (verbal reinforcement) and nonsupport (attention but no verbal reinforcement) than did Ss rated as having less deprivation in their preinstitutional histories. A follow-up study of these Ss three years later (Zigler and Williams, 1963) yielded no significant differences between Ss with histories of high and low preinstitutional deprivation; all significant effects were due to reinforcement conditions. It was concluded that additional years spent in an institution will "increase the motivation for all social reinforcers and correspondingly increase the sensitivity to differing degrees of social reinforcement (e.g., approval versus attention alone)" (p. 203). However, Stevenson and Knights (1962) tested institutionalized retardates at varying time periods after they had spent a summer vacation at home and found that the
only Ss increasing performance under conditions of social reinforcement were girls tested immediately after their return to the institution. These investigators concluded that "separation from home and family results in more intense feelings of deprivation and isolation than occur after the child has readjusted to the institutional setting" (p. 593). Butterfield and Zigler (1965) have pointed out that "institutionalization" is not a homogeneous psychological variable and is not necessarily equated with social deprivation. They compared Ss from two institutions with "markedly different social climates" and found a significant difference, with Ss from the more homelike institution spending less time on a task under conditions of social reinforcement than Ss from the less homelike institution.

The studies reviewed above emphasized the effects of contingent social reinforcement on the behavior of institutionalized mental retardates. In the current study, however, neither the contact with E in the instruction period nor the extra E-contact which occurred with 12 Ss could be described as contingent social reinforcement. In both situations the behavior of E was neutral and was directed toward instructing the S, or in the case of extra E-contact, answering a question or repairing the apparatus. The behavior of E in no way involved verbal approval of the Ss' performance. It is likely, however, that the extra
E-contact provided some sort of informational cues to the Ss, indicating to them that they should continue responding.

The remaining factors which could have had reinforcing properties which tended to equalize the effect of source of reinforcer selection were the score lights and buzzer in all groups and token reinforcement in the experimental groups. There have been other studies in which token reinforcement may have obscured differences in the effects of backup reinforcers (Bravenec and Estes, 1969; Estes, 1965; Liebert and Allen, 1967; Ryan, 1963). In addition, differences in the magnitudes of token reinforcement may have contributed to the significant differences Heber (1959) found between his high and low incentive conditions. In addition to Ss working for high and low incentives (first and seventeenth choice of reinforcers), they also received different amounts of token reinforcement, three marbles/trial in the high incentive condition and one marble/trial in the low incentive condition. Therefore, incentive conditions differed in magnitude of token reinforcement as well as in qualitative value of backup reinforcers.

In the current study Ss in the subject-selected reinforcer conditions worked for their first choice of reinforcer, while Ss in the experimenter-selected reinforcer conditions worked for the reinforcer whose value was ranked
at the median position (13th) in a pre-experimental pilot study. Therefore, Ss in these two conditions were working for items of different reinforcement values. Although the potential effectiveness of varying quality or reinforcement values of the reinforcers used has been emphasized by several writers (Bruning, 1965; Heber, 1959; Ryan, 1963) and studies have provided evidence of better performance in Ss working for reinforcers of high value (Blount, 1967; Heber, 1959; Siegel, Forman, and Williams, 1967; Siegel, Williams, and Forman, 1967; Williams, 1968), the negative finding of the current study with regard to this variable is not without precedent. There have been at least three studies (Bravenec and Estes, 1969; Estes, 1965; Ryan, 1963) in which normal children ranked a group of toys for their reinforcement values, but there were no differential effects on performance as a function of Ss' working for their most preferred toy versus least preferred toy or no backup reinforcer. These studies differed from the current one both in type of S and type of task. In the only study available which varied magnitude of reinforcement when self-reinforcement was involved, Liebert and Allen (1967) found no significant differences in the number of tokens Ss self-administered according to whether tokens could be traded for prizes or had no such exchange value.
In comparing the current study with Heber's (1959) study, several reasons can be hypothesized as to why Heber found significant differences in performance between high and low incentive conditions which were not found between the subject-selected reinforcer and experimenter-selected reinforcer conditions of the current study: (a) the differences in use of token reinforcement, which may have increased the differences between groups in Heber's study but tended to equalize the differences in the current study, (b) the reinforcers used in Heber's study had a wider range of monetary value, and (c) Heber used each S as his own control, so each S worked under both high and low incentive conditions. Gerjuoy and Winters recommended use of that approach: "Since inter-subject variability is frequently larger in retarded Ss than in normal Ss, it is usually preferable, whenever possible, to use each S as his own control" (1969, p. 6).

With regard to the reinforcer value of items used in the subject-selected reinforcer conditions of the current study, it was noted that the reinforcers chosen by Ss in these groups differed somewhat from the reinforcers selected by Ss in the pilot study for ranking of reinforcement values. In the pilot study 8 different items were selected as first choice of reinforcer by the 10 Ss, whereas only 9 different items were selected by 28 Ss in the
experimental groups. Of the items chosen four were selected by more than one S; 11 Ss chose the quarter, 5 the bar of soap, 3 the plastic ring, and 2 the comic book. The soap had the highest mean ranking in the pilot study, but the quarter was ranked twelfth. Although there were insufficient data for statistical analysis, the quarter seemed to be especially popular among male Ss age 21 and older, while the soap was valued primarily by older females. In similar studies using older institutionalized retardates as Ss, efforts should be made to provide more reinforcers which would be valued by this age group.

From the review of the literature and results of the current study, the following conclusions can be made about use of experimenter-selected versus subject-selected reinforcers. Studies have generally shown that varying magnitudes of a reinforcer specified by E produces no significant differences in performance in normals (Bruning, 1964; Miller and Estes, 1961; Weinstock, 1965) or mental retardates (Cantor and Hottel, 1955; Cromwell and Moss, 1959; Ellis, 1962). A more promising approach has been that of having Ss rank items according to their reinforcement value and defining high and low reinforcement according to Ss' rankings. While such an approach has yielded significant differences in performance between reinforcement conditions (Blount, 1967; Heber, 1959; Siegel, Forman,
and Williams, 1967; Siegel, Williams, and Forman, 1967; Williams, 1968), the current study as well as other studies (Bravenec and Estes, 1969; Estes, 1965; Ryan, 1963) found no significant differences as a function of reinforcement conditions. The current study differed from the others in that Ss' reinforcement values were used to determine only the "high" reinforcement condition, while the "low" reinforcement condition was specified by E. It is concluded that differences in performance found between reinforcement conditions determined by Ss' rankings of reinforcement values will depend upon various factors, including the following:

1. The effects of token reinforcement, whether different groups receive the same magnitude of token reinforcement (which might tend to equalize the effects of the different backup reinforcers) or different magnitudes of token reinforcement (which might tend to maximize differences between the reinforcement conditions).

2. Control over the occurrence of social reinforcement, so that its effects do not interfere with the effects of other reinforcement conditions.

3. The presence or absence of other uncontrolled factors which may have reinforcing properties (e.g., the lights and buzzer in the current study).
4. The actual magnitude of the differential reinforcement values within the reinforcer pool. While this variable would differ with each S, it would seem reasonable that a reinforcer pool with a wide range of monetary value would include greater variation in reinforcement value than a reinforcer pool which is restricted to a narrow range.

5. Whether groups under different reinforcement conditions are compared or whether each S serves as his own control, performing under each reinforcement condition.

An important point regarding the failure of the subject-selected reinforcer conditions to increase performance over that of the experimenter-selected reinforcer conditions is that while Ss selected their own reinforcer, they made their choice from a pool of items selected by E. More effective reinforcers might be obtained by letting Ss specify their own reinforcers. It would also be potentially more effective to include other types of reinforcers, such as activities and social reinforcers, rather than restricting the reinforcer selection to material items, especially when all material items are of such low monetary value. While use of such procedures would be more complex for E to deal with and some restrictions would have to be
made on Ss' choices of reinforcers, it is hypothesized that reinforcers selected by Ss in this manner would have a greater effect on behavior than reinforcers selected by Ss from a pool of items selected by E.

The factors discussed above as contributing to lack of differences in performance as a function of source of reinforcer selection also apply to the failure to find a significant difference between conditions of self-reinforcement and external reinforcement. In addition, due to deficiencies in these Ss' self-reinforcement processes and to certain characteristics of the experimental procedure, certain elements of the self-reinforcement conditions functioned more like external reinforcement. Although Ss in the self-reinforcement conditions initially selected a criterion, the control was immediately switched to E's remote panel, and the criterion remained the same regardless of Ss' subsequent manipulation of the criterion switch. Therefore, after the initial setting, the criterion was externally controlled rather than controlled by the S.

In spite of the lack of difference between the performance of Ss in self-reinforcement and external reinforcement conditions, the current study provided some dramatic evidence that institutionalized mental retardates have not developed effective processes of self-reinforcement, as
they were described by Bandura and Perloff (1967, p. 113):

There are two important elements within a self-reinforcing event whose independent effects must be assessed before persistence of self-reinforced behavior can be meaningfully interpreted. These are (a) the self-imposition of an achievement standard, and (b) the self-administration of rewards.

It was apparent from the capricious manner in which Ss in the current study changed the setting of the criterion switch that such settings did not represent the "self-imposition of an achievement standard." Gerjuoy and Winters (1969) have concluded that compared with normals retardates' criteria of success are lower. However, the results of the current study indicate that rather than simply being lower, their criteria for performance are inconsistent and not well established.

The Ss' failure to set performance standards for themselves did not appear to be a function of the instructions. These instructions complied with all of Gerjuoy and Winters' (1969) recommendations regarding instructions for mental retardates; they were repetitious and used simple, concrete language, verbal instructions were supplemented with non-verbal demonstrations, and Ss were required to make responses indicating understanding of the instructions. Ss were able to follow all aspects of the instructions except those which involved selecting and setting a performance criterion for themselves. Although Stevenson
and Cruse's (1961) study did not involve self-reinforcement, one of the results of that study is of interest here. The performance of institutionalized retardates was compared under conditions of social reinforcement, attention, and absence of E. The absence condition did not work due to Ss' failure to follow the instructions. While these Ss were younger and somewhat lower level (mean CA 14-2, mean MA 6-2) than Ss in the current study, the instructions and task were much simpler. Therefore, it can be argued that Ss' failure to follow instructions was not a function of lack of understanding but rather of an inability to maintain their own behavior in the absence of external reinforcement, i.e., to self-reinforce.

Ss in the current study were able to self-administer token reinforcement. As hypothesized, Ss self-administered significantly greater magnitudes of token reinforcement than did Ss in Bandura and Perloff's (1967) study. Magnitude of self-reinforcement was not a function of either sex or the criterion originally set. On the basis of these findings it can be concluded that institutionalized retardates are more generous than normal children in self-reinforcement but that self-reinforcement in mental retardates is not contingent upon performance. These results provide an interesting comparison with the findings of Masters (1968, 1969) and Kanfer (1966). In
Masters' studies. Ss who had less success (received less reinforcement) than their partners indulged in subsequent generous noncontingent self-reinforcement. Kanfer found that Ss who were ranked in the lower half of their classes engaged in more inappropriate self-reinforcement than did Ss ranking in the upper half of their classes. The teachers' ratings correlated significantly with IQ, indicating that Ss with lower IQs tended to self-reinforce after an incorrect response more often than Ss with higher IQs. The self-reinforcement in the current study could not be judged appropriate or inappropriate, as the responses could not be categorized as being correct or incorrect. However, the self-reinforcement of Ss in the current study was noncontingent (at least upon a performance standard set by the Ss themselves); therefore such generous noncontingent self-reinforcement might be considered to be a function of the cumulative failures and inadequacies that these institutionalized retardates have experienced in comparison with their peers. It is also possible that the patterns of self-reinforcement exhibited in the current study were partially a function of the absence of a model or specific instructions for dealing with this unfamiliar task. The patterns of self-administration of token reinforcement in the current study were comparable to those exhibited by Ss in Bandura and Kupers' (1964) control group who had not been
exposed to a model; in this group "self-rewards were apparently freely dispensed and not made contingent on meeting or surpassing any minimum standard of achievement" (p. 8).

The hypothesis that males would make significantly more responses than females was based on the fact that Bandura and Perloff (1967) found such a difference between sexes. The difference was attributed to "the fact that the task required some physical effort, and consequently the boys' higher output reflects their greater strength" (p. 115). The finding of no significant performance differences between the sexes in the current study raises the questions of whether the sex difference found in Bandura and Perloff's study was actually attributable to the factor of strength and whether their finding was specific to normal children and/or to the particular age group used.

The final hypothesis made in the current study was that performance of the control group would be lower than that of the experimental groups. With the sexes in the control group combined, there were no significant differences. However, there was a significant difference between the sexes in the control group. Separate comparisons of males and females in the control group with the experimental groups revealed that females in the control group had significantly lower scores than the experimental groups, as hypothesized, but that there were no significant differences between performance scores of the experimental groups.
and males in the control group. While this sex difference in the control group was unexpected, somewhat similar phenomena have occurred in other studies. With institutionalized retardates, Stevenson and Knights (1962) found that boys made more (but not significantly more) responses under a neutral condition than under a condition of social reinforcement. To account for this finding, it was hypothesized that "the sex of E has a differential effect on the performance of boys and girls, such that social reinforcement delivered by a male E is more effective in modifying the performance of girls than of boys" (p. 593). Such an explanation cannot be applied to the results of the current study, however, since E was female and the reinforcement involved was not primarily social. With normal kindergarten children divided into high and low persistence groups, Nakamura and Ellis (1964) found a high level of activity in the low reward-low persistence group for which they had no explanation. With third graders they found significant sex differences, including increased activity during extinction on the part of the boys. These finding are similar to the high level of responding of control group males in the current study.

It would seem questionable to attribute the difference between sexes in the control group of the current study to greater strength or endurance on the part of male
Ss when no such sex difference was found in the experimen-
tal groups. The sources of reinforcement that were avail-
able to control group Ss could be considered--including the
lights and buzzer for achievement of criterion, the contact
with E, and even the possibility that the reinforcement was
not perceived of as noncontingent upon performance--but
these possible sources of reinforcement were equally appli-
cable to males and females and therefore do not account for
the sex difference that occurred. One feature of the con-
trol group which did not characterize the experimental
groups was the fact that no pause in responding was re-
quired for receipt of reinforcement when a criterion was
reached. Therefore, Ss in the control group could keep re-
sponding without breaks and thereby emit more responses in
an equivalent time period than Ss in the experimental
groups. This factor, plus an orientation on the part of
the males toward physical activity, might account for the
difference. On the Performance Recording Sheets of the
control group males, it was noted that three of the seven
Ss turned the crank at an extremely fast rate with few
pauses. In response to the Experimental Demand Questions,
one S even reported that he thought he was supposed to
"keep it going steady and ring it quicker than if you go
slow." In contrast, none of the females in the control
group turned the crank at a particularly fast rate.
Perhaps the issues of noncontingent reinforcement and effects of instructions need to be explored further. Evans and Spradlin's (1966) study, which was described in detail in the Background Chapter, suggested that Ss may view reinforcement as being contingent upon their performance even when that is not actually the case. While it might be hypothesized that this was the case with males in the control group in the current study, there is no evidence to support this conclusion and it would be difficult to explain why this should be the case for males but not for females.

While instructions have been shown to have important effects on the performance of mental retardates (Evans and Spradlin, 1966; Headrick, 1963; Stevenson and Zigler, 1957), the effects of instructions in the current study are difficult to evaluate, as instructions were not systematically varied. Certainly the instructions were not explicit with regard to the criteria and magnitude of self-reinforcement; these decisions were left to the Ss. The results indicated that these Ss were unable to function effectively when the choice of criterion was left up to them. The results obtained by Bandura and Perloff using similar instructions with normal children were quite different. Since most criteria for self-reinforcement are determined in the absence of objective criteria for
adequate performance, it would be worthwhile to study the behavior of various types and ages of Ss under conditions in which no explicit rules were given for self-reinforcement. Apart from Bandura and Perloff's study, the only available information on this issue comes from data on control groups in various studies. Since Ss, tasks, and other experimental variables differ so much in these studies, it is difficult to derive specific conclusions from a comparison of their results.

It is difficult to account for the large variance within groups found in the current study. Gerjuoy and Winters stated that "inter-subject variability is frequently larger in retarded Ss than in normal Ss" (1969, p. 6), but they did not speculate as to why this might be the case. One of the factors contributing to the high variance within groups in the current study was the additional contact with E experienced by some Ss subsequent to the instruction period. The fact that these Ss made significantly more responses than Ss with no extra E-contact testifies to the influence of this variable on performance. The failure on the part of E to eliminate Ss having extra E-contact was unfortunate. While it was established by statistical analysis that extra E-contact was an important influence on performance, no adequate evaluation could be made of the probable results of the study had these Ss
been eliminated at the time that the deviation from standard experimental procedure occurred. Eliminating the data on these Ss and those yoked to them and analyzing the data on remaining Ss provided inconclusive results, partially because of the small size of the remaining sample (N = 32) and the unequal group ns.

It is possible that some of the within-group variance found might be attributable to factors not considered in the current study. Two such factors are length of institutionalization and general persistence level. Zigler and Williams (1963) found that increased length of institutionalization obscured the effects of amount of preinstitutional deprivation and increased sensitivity to variations in social reinforcement. It might well be that length of institutionalization is also an important variable in determining the effects of other types of reinforcement, including self-reinforcement. In addition to length of institutionalization, there might be differences in general behavioral characteristics which were not identified or varied systematically in the current study but which could account for some of the within-group variance in performance scores. One such behavioral characteristic, level of persistence, was considered by Nakamura and Ellis (1964) and Nakamura and Lowenkron (1964). It would be of considerable interest to determine whether discriminations could be made
among institutionalized retardates on such general behavioral characteristics as level of persistence, which are a function of their previous reinforcement histories both before and after entering the institution, and whether differences in these behavioral characteristics were reflected in differences in performance.

Another characteristic according to which Ss have been classified, which is of special relevance to the current study, is that of extrinsic versus intrinsic motivation. Milgram (1969) has summarized Zigler's conclusions that "retardates tend to be outer-directed rather than inner-directed in their problem solving ability, looking to external cues rather than to their own frames of reference or fund of information for solutions to problems" (p. 533). The study by Turnure and Zigler (1964) illustrated this point. Retardates were found to be more imitative than normals regardless of whether such imitation facilitated or hindered performance. Turnure and Zigler attributed retardates' use of an outer-directed approach to problem solving to their history of failure experiences. Haywood and Weaver (1967) found incentive effects to be dependent upon extrinsic versus intrinsic motivation patterns, but reported that it was difficult to find Ss who were relatively intrinsically motivated among institutionalized retardates. If self-reinforcement is considered to be
analogous to intrinsic motivation patterns and external reinforcement analogous to extrinsic motivation patterns, then these findings are in accord with those of the current study—that the majority of institutionalized retardates have not developed self-reinforcement or intrinsic motivation patterns.

Experimental Demand Questions were asked of all Ss in an attempt to determine how they viewed the experimental task. In none of the previous studies of self-reinforcement has there been a means of evaluating the Ss' expectations and conclusions about the task. Rather, it has been assumed from the results of the studies that a self-reinforcement process has been operative. Consequently it is often difficult to determine whether Ss are actually setting standards for their own behavior and reinforcing themselves according to achievement of these standards or whether they are responding to implicit messages to imitate a model or to behave according to a certain aspect of the instructions.

When the Experimental Demand Questions were constructed, it was the opinion of E that the institutionalized retardates who served as Ss might be able to conform to the instructions given for the self-reinforcement conditions, yet be unable to conceptualize and/or verbalize the self-reinforcement process. The results indicated
that they did neither successfully. The deficits in self-reinforcement exhibited by these Ss have already been discussed. Responses to the Experimental Demand Questions were described in the Results Chapter; the majority of the responses were random, concrete, vague, or "I don't know" responses. No S gave a response incorporating the idea of setting a performance standard and self-reinforcing when that standard was achieved.

These results are not surprising when results of previous studies are considered. Klugh and Janssen (1966) reviewed several studies which indicated "a deficit in the retardate's ability to verbalize the characteristics of the correct discriminanda once he has achieved success by a trials correct criterion" (p. 903). Such was the case in Klugh and Janssen's own study; among institutionalized adult retardates who had discriminated a square as being the correct stimulus on ten consecutive trials, only 41% were able to verbalize the rule for selecting the correct stimulus for both simultaneous and successive presentations and only 14% were able to do so for simultaneous presentations. Some of the mentally retarded Ss who had reached criterion verbalized an incorrect concept and made subsequent erroneous discrimination choices on the basis of the incorrect concept that they had verbalized. "Within a few trials the errors ceased and S would achieve the
consecutively correct responses, whereupon he would often give an incorrect rule again" (p. 906). This occurred even though retarded Ss reached criterion as frequently as normal Ss of the same MA who were "much more likely to verbalize correctly" (p. 906). Verbalization of relevant cues or of experimental demand characteristics appears to be a rather high-level skill. With normal nine-year-old children Miller and Estes (1961) found that less than half of the Ss who learned to select the correct stimulus in a discrimination task were able to verbalize the relevant cue, but there were no significant differences in number of errors or trials to criterion between Ss who could identify the relevant cue and those who could not. Compared with the above studies, Ss in the current study were asked much more complex questions. It is not surprising, therefore, that they were unable to give adequate responses to these questions and in doing so to verbalize their concepts of the purpose and demands of the experiment. Administering experimental demand questions to older, intellectually normal Ss might be expected to yield much more useful information about Ss' perceptions of the purpose and demands of an experiment.

The findings with the most far-reaching implications in the current study were that institutionalized mental retardates do not set consistent standards for their
own behavior and that when they self-administer rein-
forcement it is generous but not contingent upon their
performance. This lack of established standards for self-
reinforcement is likely a function of multiple variables,
including the absence of consistent external standards
applied to their behavior in the institution, previous ex-
periences of failure, and lack of effective models in the
institutional environment. Studies of the effects of mod-
els have shown that adults exert a more powerful modeling
influence than peers (Bandura and Kupers, 1964), but that
the presence of conflicting peer standards decreases an
adult model's influence (Bandura, et al., 1967) as do dis-
crepancies in self-reinforcement criteria exhibited by two
models (McMains and Liebert, 1968) and discrepancies be-
tween the observed and imposed reinforcement criteria
(Mischel and Liebert, 1966). Also, children may reject the
self-reinforcement standards of a model whose abilities are
too discrepant from their own (Bandura and Whalen, 1966).
In applying these experimental findings to the institution-
al environment, it can be concluded that the understaffed
conditions of most institutions provide few desirable adult
models, and the influence of these models may be diminished
by the discrepancies between their capabilities and those
of the mental retardates, discrepancies in self-
reinforcement criteria exhibited by various models,
discrepancies between modeled and imposed criteria, and the influence of conflicting peer standards.

Bandura has stated that "persons who have failed to develop self-monitoring reinforcement systems . . . require considerable social surveillance to ensure that they do not transgress" (1969, pp. 37-38). While this statement referred primarily to antisocial persons such as juvenile delinquents, it applies equally well to institutionalized mental retardates. If the staffs of institutions could be used effectively to train retardates in self-reinforcement processes, the amount of supervision required in these institutions could be decreased. Bandura (1969) and Marston (1965b) have both indicated that such training is possible and desirable. Bandura stated that self-reinforcement processes are "socially transferable and conditionable" (1969, p. 621) and are "essential if induced behavioral changes are to generalize and endure, particularly where social environments provide either weak support for new modes of behavior or conflicting patterns of reinforcement" (1969, p. 624). Marston (1965b) has discussed the role of the therapist in improving patients' accuracy of self-evaluation and increasing their reliance on positive self-reinforcement. There is no obvious reason why the procedure Bandura described for training antisocial persons in self-reinforcement could not be applied to institutionalized retardates (1969, p. 620):
This is achieved by gradually transferring evaluative and reinforcing functions from change agents to the individual himself. Rewards are now made contingent not only upon occurrence of desired behavior, but also accurate evaluation of one's own performances. Although at this stage the person judges when his behavior warrants reward according to the prevailing contingency structures, others still serve as the reinforcing agents. After accurate self-evaluative behavior is well established, the reinforcing function is likewise transferred so that the individual both evaluates his own behavior and reinforces himself accordingly. In addition, the artificial material rewards are gradually reduced as the person's behavior is brought increasingly under the control of self-administered and symbolic consequences. The ultimate aim of training in self-reinforcement is to produce a level of functioning at which participants can control their own behavior with minimum external constraints and artificial inducements.

Bandura also recommended having such persons serve as evaluative and reinforcing agents for their peers' behavior, assuming that they will come to apply the same standards to their own behavior and emphasizing that social reinforcement is required to prevent extinction of self-reinforcing behaviors. Therefore, the best approach to developing self-reinforcing standards and behaviors would be to use groups of peers who could reinforce each other as well as themselves for attaining the specified standards.

While it is likely that the lack of self-prescribed standards for self-reinforcement in institutionalized mental retardates is attributable to deficits in their environment, it would be important to conduct research on self-reinforcement in noninstitutionalized mental
retardates. Comparison of the results of such research with the results of the current study would help delineate the relative importance of the various factors which contribute to the development of self-reinforcement processes. If noninstitutionalized mental retardates were found to have developed effective self-reinforcement processes, such a finding would emphasize the importance of the environmental deficits in institutionalized retardates' failure to develop self-reinforcement. If noninstitutionalized retardates were found to have deficits in self-reinforcement similar to those of institutionalized retardates, then experiences common to all retardates, such as frequent failure, could be considered to be crucial in the lack of development of self-reinforcement procedures. Such comparative data would be valuable in determining where emphasis should be placed in training institutionalized retardates in self-reinforcement processes and in determining whether noninstitutionalized retardates are in need of similar training programs in self-reinforcement.
APPENDIX A

CA AND IQ DATA ON ELIGIBLE SUBJECT POOL AND EXPERIMENTAL SS

Table 4. CA Distributions according to Sex of the 116 Eligible Ss

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Number and Percentage of Ss</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-14</td>
<td></td>
<td>10 (14.71%)</td>
<td>3 (06.25%)</td>
</tr>
<tr>
<td>15-19</td>
<td></td>
<td>36 (52.94%)</td>
<td>17 (35.42%)</td>
</tr>
<tr>
<td>20-24</td>
<td></td>
<td>9 (13.23%)</td>
<td>9 (18.75%)</td>
</tr>
<tr>
<td>25-29</td>
<td></td>
<td>11 (16.18%)</td>
<td>8 (16.67%)</td>
</tr>
<tr>
<td>30-34</td>
<td></td>
<td>2 (02.94%)</td>
<td>11 (22.92%)</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>68 (100.00%)</td>
<td>48 (100.01%)</td>
</tr>
</tbody>
</table>
Table 5. IQ Distributions according to Sex of the 116 Eligible Ss

<table>
<thead>
<tr>
<th>IQ Range</th>
<th>Number and Percentage of Ss Male</th>
<th>Number and Percentage of Ss Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-54</td>
<td>11 (16.18%)</td>
<td>9 (18.75%)</td>
</tr>
<tr>
<td>55-59</td>
<td>21 (30.88%)</td>
<td>7 (14.58%)</td>
</tr>
<tr>
<td>60-64</td>
<td>20 (29.41%)</td>
<td>21 (43.75%)</td>
</tr>
<tr>
<td>65-70</td>
<td>16 (23.53%)</td>
<td>11 (22.92%)</td>
</tr>
<tr>
<td>Totals</td>
<td>68 (100.00%)</td>
<td>48 (100.00%)</td>
</tr>
</tbody>
</table>
Table 6. Mean CAs according to Sex of Experimental Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Self-reinforcement, subject-selected reinforcer group</td>
<td>22-7.71</td>
<td>22-7.86</td>
</tr>
<tr>
<td>B: Self-reinforcement, experimenter-selected reinforcer group</td>
<td>22-8.57</td>
<td>22-9.00</td>
</tr>
<tr>
<td>C: External reinforcement, subject-selected reinforcer group</td>
<td>22-8.57</td>
<td>22-7.57</td>
</tr>
<tr>
<td>D: External reinforcement, experimenter-selected reinforcer group</td>
<td>22-7.57</td>
<td>22-8.71</td>
</tr>
<tr>
<td>E: Noncontingent reinforcement control group</td>
<td>22-7.14</td>
<td>22-8.86</td>
</tr>
</tbody>
</table>

Table 7. Mean IQs according to Sex of Experimental Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Self-reinforcement subject-selected reinforcer group</td>
<td>61.14</td>
<td>61.43</td>
</tr>
<tr>
<td>B: Self-reinforcement, experimenter-selected reinforcer group</td>
<td>60.57</td>
<td>60.71</td>
</tr>
<tr>
<td>C: External reinforcement, subject-selected reinforcer group</td>
<td>60.71</td>
<td>61.00</td>
</tr>
<tr>
<td>D: External reinforcement, experimenter-selected reinforcer group</td>
<td>60.43</td>
<td>61.00</td>
</tr>
<tr>
<td>E: Noncontingent reinforcement control group</td>
<td>60.86</td>
<td>61.29</td>
</tr>
</tbody>
</table>
APPENDIX B

APPARATUS DIAGRAMS
Figure 1. Front Panel of the Apparatus
Figure 2. Experimenter's Remote Control Panel
### APPENDIX C

#### RESULTS OF PILOT STUDY A

Table 8. Mean Reinforcer Rankings for Five Male and Five Female Pilot Study Ss Combined

<table>
<thead>
<tr>
<th>Reinforcer</th>
<th>Mean Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soap</td>
<td>8.6</td>
</tr>
<tr>
<td>Comic book</td>
<td>8.8</td>
</tr>
<tr>
<td>Soap bubbles</td>
<td>9.0</td>
</tr>
<tr>
<td>Coloring book</td>
<td>9.2</td>
</tr>
<tr>
<td>Gum</td>
<td>9.9</td>
</tr>
<tr>
<td>Playing cards</td>
<td>10.2</td>
</tr>
<tr>
<td>Crayons</td>
<td>10.9</td>
</tr>
<tr>
<td>Emery boards</td>
<td>11.7</td>
</tr>
<tr>
<td>Perfume</td>
<td>11.9</td>
</tr>
<tr>
<td>Water pistol</td>
<td>12.0</td>
</tr>
<tr>
<td>Comb</td>
<td>12.3</td>
</tr>
<tr>
<td>Quarter (25¢)</td>
<td>12.4</td>
</tr>
<tr>
<td>Fritos</td>
<td>12.7</td>
</tr>
<tr>
<td>Cupcakes</td>
<td>12.9</td>
</tr>
<tr>
<td>Toy watch</td>
<td>14.1</td>
</tr>
<tr>
<td>M &amp; Ms</td>
<td>15.2</td>
</tr>
<tr>
<td>Cracker Jacks</td>
<td>15.4</td>
</tr>
<tr>
<td>Balloon</td>
<td>15.4</td>
</tr>
<tr>
<td>Gumdrops</td>
<td>15.6</td>
</tr>
<tr>
<td>Peanuts</td>
<td>15.7</td>
</tr>
<tr>
<td>Barrettes</td>
<td>16.9</td>
</tr>
<tr>
<td>Ring</td>
<td>16.9</td>
</tr>
<tr>
<td>Whistle</td>
<td>17.0</td>
</tr>
<tr>
<td>Candy bar</td>
<td>17.2</td>
</tr>
<tr>
<td>Hilo game</td>
<td>18.1</td>
</tr>
<tr>
<td>Truck</td>
<td>18.9</td>
</tr>
</tbody>
</table>
APPENDIX D

PERFORMANCE RECORDING SHEETS AND EXPERIMENTAL DEMAND QUESTIONS
Performance Recording Sheet:
Self-Reinforcement Conditions

Subject:
CA:
IQ:
Experimental Group:

Performance
Performance criterion set at ____ points.
Total number of crank revolutions (from counter): ____
Total number of completed runs (achieving criterion and rewarding self): ____
Total number of tokens taken: ____
Average number of tokens taken/reward period: ____
Performance Recording Sheet:
External Reinforcement Conditions

Subject:
CA:
IQ:
Experimental Group:

Yoked to:
Subject:
CA:
IQ:
Experimental Group:
Performance criterion set at _____ points.
Average number of tokens taken/reward period: _____
(Total number of tokens taken: _____)
(Total number of crank revolutions: _____)

Performance
Total number of crank revolutions (from counter): _____
Total number of completed runs (achieving criterion and being rewarded): _____
Total number of tokens dispensed: _____
Performance Recording Sheet:

Noncontingent Reinforcement Control Condition

Subject:
CA:
IQ:

Yoked to:
Subject:
CA:
IQ:

Experimental Group: B (Self-reinforcement, experimenter-selected reinforcer condition)

Performance criterion set at ____ points.
Total number of tokens taken: ____
(Total number of crank revolutions): ____

Performance
Total number of crank revolutions: ____
Experimental Demand Questions:
Self-Reinforcement Conditions

Subject: IQ:
CA: Experimental Group:

These questions are to be asked of each S after he has stopped performing and has been given his backup reinforcer. Questioning will begin with the following open-ended questions:

1. What do you think this study is about?
2. What do you think you were supposed to do?

An attempt will then be made to clarify responses to the above questions and to obtain omitted information by means of the following specific questions:

3. Why did you stop turning the crank when you did?
4. How long did you think you were supposed to turn it?
5. How many tokens did you think you needed to get the prize?
6. Did you think you had earned enough tokens to get the prize?
7. Why did you set the machine at _____ points?
8. Where did you think you were supposed to set it—at how many points?
9. Why did you take _____ tokens each time?
   Or: Why did you take _____ tokens sometimes and _____ tokens other times?
10. How many tokens did you think you were supposed to take each time?
Experimental Demand Questions:

External Reinforcement Conditions

Subject: IQ:
CA: Experimental Group:

These questions are to be asked of each S after he has stopped performing and has been given his backup reinforcer. Questioning will begin with the following open-ended questions:

1. What do you think this study is about?

2. What do you think you were supposed to do?

An attempt will then be made to clarify responses to the above questions and to obtain omitted information by means of the following specific questions:

3. Why did you stop turning the crank when you did?

4. How long did you think you were supposed to turn it?

5. How many tokens did you think you needed to get the prize?

6. Did you think you had earned enough tokens to get the prize?
Experimental Demand Questions:

Noncontingent Reinforcement Control Condition

Subject:

CA:

IQ:

These questions are to be asked of each S after he has stopped performing and has been given his backup (non-contingent) reinforcer. Questioning will begin with the following open-ended questions:

1. What do you think this study is about?
2. What do you think you were supposed to do?

An attempt will then be made to clarify responses to the above questions and to obtain omitted information by means of the following specific questions:

3. Why did you stop turning the crank when you did?
4. How long did you think you were supposed to turn it?

Andronico, M. P. The effects of different reinforcing cues upon the free operant response rates of mental retardates and normal children of equated CA or MA. Dissertation Abstracts, 1964, 24(7), 2981-2982. (Abstract)


Marston, A. R. Self-reinforcement: The relevance of a concept in analogue research to psychotherapy. Psychotherapy: Theory, Research & Practice, 1965, 2, 1-5. (b)


Watson, L. S. Application of operant conditioning tech­
niques to institutionalized severely and profoundly
retarded children. Mental Retardation Abstracts,

Watson, L. S., and Lawson, R. Instrumental learning in
mental retardates. Mental Retardation Abstracts,

Watson, L. S., Orser, R., and Sanders, C. Reinforcement
preferences of severely mentally retarded children
in a generalized reinforcement context. American
Journal of Mental Deficiency, 1968, 72, 748-756.

Weinstock, A. C. The effects of magnitude of reward on
discrimination learning in preschool children.
Unpublished master's thesis, State University of
Iowa, 1962. Cited by Ryan, T. J., The effects of
nonreinforcement and incentive value on response

White, G. M. The effect of type of reward on learning in
educable mentally retarded pupils. Dissertation
Abstracts, 1968, 29(7-A), 2607. (Abstract)

Williams, J. Reversal learning in retardates as a function
of overlearning and incentive preference value.
Dissertation Abstracts, 1968, 29(2-B), 799-800.
(Express)

Witryol, S. L., and Fischer, W. F. Scaling children's in­
centives by the method of paired comparisons. Psy­
chological Reports, 1960, 7, 471-474.

Wolfensberger, W. Differential rewards as motivating fac­
tors in mental deficiency research. American Jour­
nal of Mental Deficiency, 1960, 64, 902-906.

Zigler, E. Social deprivation and rigidity in the perform­
ance of feebleminded children. Journal of Abnormal

Zigler, E., Hodgden, L., and Stevenson, H. The effect of
support on the performance of normal and feeble­
minded children. Journal of Personality, 1958, 26,
106-122.

Zigler, E., and Williams, J. Institutionalization and the
effectiveness of social reinforcement: A three-year
follow-up study. Journal of Abnormal and Social