INFORMATION TO USERS

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

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

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

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

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

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

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

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
EVANS, David Francis, 1948-
EFFECTS OF BEHAVIOR MANAGEMENT TRAINING ON
TEACHERS' ATTENTION AND THE BEHAVIOR OF
THEIR RETARDED PUPILS.

The University of Arizona, Ph.D., 1975
Psychology, clinical

Xerox University Microfilms, Ann Arbor, Michigan 48106

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.
EFFECTS OF BEHAVIOR MANAGEMENT TRAINING ON TEACHERS' ATTENTION AND THE BEHAVIOR OF THEIR RETARDED PUPILS

by

David Francis Evans

A Dissertation Presented 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

1975
I hereby recommend that this dissertation prepared under my direction by David Francis Evans entitled EFFECTS OF BEHAVIOR MANAGEMENT TRAINING ON TEACHERS' ATTENTION AND THE BEHAVIOR OF THEIR RETARDED PUPILS be accepted as fulfilling the dissertation requirement of the degree of Doctor of Philosophy.

[Signature]
Dissertation Director

[Signature] [Date]

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

[Signature] [Date]

[Signature] [Date]

[Signature] [Date]

[Signature] [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: [Signature]
ACKNOWLEDGMENTS

The author expresses his gratitude to the following people who helped so much in the completion of this investigation: Phil Balch, Gene Evans, Mary Itani, Paula Stone-Bender, and the teachers at Ganoung School.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>4</td>
</tr>
<tr>
<td>Teacher Training Manuals and</td>
<td>4</td>
</tr>
<tr>
<td>Programmed Instruction Texts</td>
<td></td>
</tr>
<tr>
<td>Direct Training of Teachers by</td>
<td>5</td>
</tr>
<tr>
<td>Mental Health Professionals</td>
<td></td>
</tr>
<tr>
<td>Single Subject Interventions</td>
<td>6</td>
</tr>
<tr>
<td>Single Subject and Classroom Interventions with Increased Teacher Participation</td>
<td>7</td>
</tr>
<tr>
<td>Comprehensive Teacher Training Programs</td>
<td>12</td>
</tr>
<tr>
<td>3. PURPOSE AND CONTEXT OF THE PRESENT STUDY</td>
<td>17</td>
</tr>
<tr>
<td>4. METHOD</td>
<td>19</td>
</tr>
<tr>
<td>Subjects</td>
<td>19</td>
</tr>
<tr>
<td>General Procedures</td>
<td>20</td>
</tr>
<tr>
<td>Data Collection</td>
<td>21</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>24</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>25</td>
</tr>
<tr>
<td>5. RESULTS</td>
<td>27</td>
</tr>
<tr>
<td>6. DISCUSSION</td>
<td>51</td>
</tr>
<tr>
<td>APPENDIX A: DEFINITIONS OF CODING CATEGORIES</td>
<td>64</td>
</tr>
<tr>
<td>APPENDIX B: SUMMARY OF TRAINING PROCEDURES</td>
<td>67</td>
</tr>
<tr>
<td>APPENDIX C: TEACHER QUESTIONNAIRE</td>
<td>71</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>76</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reliability of Observational Measures</td>
<td>23</td>
</tr>
<tr>
<td>2. Summary of Analysis of Observational Data</td>
<td>28</td>
</tr>
<tr>
<td>3. Summary of ANOVA, Verbal Praise of Targeted Children</td>
<td>29</td>
</tr>
<tr>
<td>4. Summary of ANOVA, Verbal-neutral Statements to Targeted Children</td>
<td>32</td>
</tr>
<tr>
<td>5. Summary of ANOVA, Appropriate Verbal Behavior of Targeted Children</td>
<td>34</td>
</tr>
<tr>
<td>6. Summary of ANOVA, Appropriate Verbal Behavior of Non-targeted Children</td>
<td>36</td>
</tr>
<tr>
<td>7. Summary of ANOVA, Inappropriate Physical Behavior of Non-targeted Children</td>
<td>38</td>
</tr>
<tr>
<td>8. Summary of ANOVA, Questionnaire Scores</td>
<td>41</td>
</tr>
<tr>
<td>9. Correlations between Observational Measures, Behaviors of Teachers and Targeted Children</td>
<td>49</td>
</tr>
<tr>
<td>10. Correlations between Observational Measures, Behaviors of Teachers and Non-targeted Children</td>
<td>50</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Teachers' Verbal Praise of Targeted Children</td>
</tr>
<tr>
<td>2.</td>
<td>Teachers' Verbal-neutral Statements to Targeted Children</td>
</tr>
<tr>
<td>3.</td>
<td>Appropriate Verbal Behavior of Targeted Children</td>
</tr>
<tr>
<td>4.</td>
<td>Appropriate Verbal Behavior of Non-targeted Children</td>
</tr>
<tr>
<td>5.</td>
<td>Inappropriate Verbal Behavior of Non-targeted Children</td>
</tr>
<tr>
<td>6.</td>
<td>David's Use of the &quot;Naughty Finger&quot;</td>
</tr>
<tr>
<td>7.</td>
<td>Vera: Following Directions</td>
</tr>
<tr>
<td>8.</td>
<td>Vera: Twisting of Nose and Hands in Mouth</td>
</tr>
<tr>
<td>9.</td>
<td>Handraises during Writing Period</td>
</tr>
</tbody>
</table>
ABSTRACT

Teachers of mentally retarded elementary school children were given eight weeks of training in classroom behavior management. The first half of training was purely didactic; the second half of training included in-class practice of behavior modification skills by the teachers. Topics covered during the weekly, 90-minute sessions included: reinforcement, extinction, timeout, shaping, fading, prompting, punishment, and problem solving. Each week teachers received handouts summarizing material presented during the session. In addition to lecturing to the teachers, the trainers presented skits demonstrating various behavior modification techniques.

Pre- and posttraining measures of teachers' knowledge of behavior modification principles and techniques were taken. Weekly observations were made of teacher-pupil interaction in the classrooms. Teacher attention was coded as Verbal Praise, Verbal Reprimand, Verbal-neutral, Physical Praise, Physical Punishment, and Physical-neutral. Antecedent pupil behavior was coded as Appropriate Verbal, Inappropriate Verbal, Appropriate Physical, and Inappropriate Physical. Interactions between teachers and targeted behavior-problem pupils were coded separately from interactions between teachers and non-targeted pupils.

Of the 20 observational measures, only Appropriate Verbal behavior of targeted children showed a significant effect for didactic
training versus didactic training plus in-class practice of behavior modification skills. Inappropriate Physical behavior of pupils not designated as behavior-problem pupils showed a significant decrease over training, but was not related to type of training. Questionnaire scores showed a significant learning effect over the course of training. Several teachers reported data indicating successful interventions based on techniques learned during training.

The preponderance of negative results are discussed in the context of training methodologies. Most training programs rely heavily on the programming of stimuli that are antecedents of hoped-for changes in teachers' behavior modification skills. The author suggests that future teacher training programs be conducted along operant lines, and that teachers be reinforced for learning techniques of classroom behavior modification. Possible forms of teacher reinforcement (money, college credit, peer approval) are discussed.
CHAPTER 1

INTRODUCTION

Traditionally, mentally retarded people in the United States have been looked upon by layman and professional alike, as defective persons who cannot benefit from education and training. Up until the late 1940's, the prescribed "treatment" for the retarded was usually custodial care (and little else) in special institutions which provided little in the way of habilitative services, and which actually fostered or accelerated the mental and social retardation of their residents. The concept of mental retardation as a continuum of deficit in adaptive functioning was not widely accepted; the terms "mentally retarded" and "vegetative idiot" were generally held to mean the same thing. The mentally retarded were considered to be an unfortunate few who, for their own good, should simply be "cared for" in isolation from the rest of society (Thompson 1972).

During World War II, the scope of the mental retardation problem in this country became evident when approximately 716,000 men were rejected from the armed services on the grounds of mental deficiency. Many of these men did not conform to the popular stereotype of the mentally retarded, but were functioning at a level that would now be classified as "mild" or "moderate" retardation. This fact alone brought national attention to the problem of mental subnormality, and helped to spur an interest in discovering possible training techniques for the mentally retarded.
The first successful application of operant conditioning to the behavior of a retarded person was reported by Fuller (1949). Using sugared milk as a reinforcer, Fuller was able to operantly condition a simple motor response of a profoundly retarded teen-ager. By the middle 1960's, there was a growing body of literature that demonstrated the efficacy of operant procedures in modifying the behavior of institutionalized retarded persons (Ellis and Pryer 1958, Barrett and Lindsley 1962, Lindsley 1964, Girardeau and Spradlin 1964). The development of operant training techniques for modifying the behavior of institutionalized retarded persons has been paralleled by the growing use of operant techniques with nonretarded and less severely retarded persons in non-institutional settings. A recent development in this area is the training of teachers to modify the behavior of their pupils, in the classroom, using operant techniques (Fargo, Behrens and Nolen 1970; Brown 1971, Graziano 1971, Patterson 1971).

The literature cited below deals with methods for training teachers to use operant techniques in the classroom. Publications and research cited deal with the training of teachers of both retarded and nonretarded children. There are two reasons for including research on the nonretarded in this review. First, the methods of operant conditioning that a teacher must master in order to modify the behavior of a pupil are the same for retarded and nonretarded children alike. Within the operant conditioning paradigm, a response is a response, whether emitted by a nonretarded child or a retarded one. Secondly, the training program on which the present research is based (Stone et al. 1973)
was originally designed for training teachers of nonretarded children. The present study represents the first application of this training program to teachers of the retarded. Thus, the methodology and results of the present study are most properly presented in the context of previous research in training teachers of both retarded and nonretarded children.
CHAPTER 2

LITERATURE REVIEW

The literature on training teachers to modify the classroom behavior of their pupils reflects two rather distinct approaches to training. The first approach is that of the programmed text or training manual, which offers printed instruction in how to learn and use operant techniques. The second, more researched approach, is the direct training of teachers by psychologists and other mental health professionals familiar with operant techniques as applied to the behavior problems of children.

Teacher Training Manuals and Programmed Instruction Texts

Numerous manuals and programmed texts offering self-instruction in classroom behavior modification are available to teachers of non-retarded children (Becker, Thomas and Carnine 1969; Homme 1970, Buckley and Walker 1971, Kunzleman 1970, Sattler and Swope 1970, Valett 1970), as well as for teachers of the retarded (Bensberg 1965, Hamilton 1971, Linford, Hipsher and Silikovitz 1972; Watson 1972). These manuals emphasize basic contingency management, the use of reinforcers and punishers, shaping, prompting, fading and so on. Manuals intended for use by teachers of the nonretarded tend to emphasize classroom management of problem behavior such as talking out and fighting, and the development of school-related skills such as attention to task and
study behaviors. Manuals written for teachers of the retarded tend to also include sections on teaching self-help skills (feeding and toileting), daily living skills such as cooking and money management, and recreational skills—all of which the nonretarded child tends to learn without specific programming.

Interventions suggested in the texts cited above are based on behavior modification techniques of proven utility. Commercial success of such behavioral cookbooks is apparent from their numbers. The adequacy of such texts as tools for giving teachers the skills necessary to modify behavior in the classroom, however, has not been extensively researched. At this time, there are only two available investigations of the effects of simply reading self-instructional texts on teachers' acquisition of behavior management skills. In both studies (Walker and Buckley 1972, Golladay 1973), the authors concluded that when used in isolation, the training manuals were not effective tools for training teachers to use behavior management skills in the classroom. Whether certain teachers can acquire behavior management skills by reading a programmed text or training manual, however, remains an empirical question. Characteristics of manuals and of teachers who attempt to learn from them are in need of further investigation.

*Direct Training of Teachers by Mental Health Professionals*

The literature reviewed below deals with efforts to train teachers as behavioral interventionists in their own classrooms. It is divided into three sections: (1) single subject interventions with only minimal teacher training; (2) single subject and classroom-wide
interventions with greater teacher involvement in training and/or intervention, and; (3) more comprehensive teacher training programs.

Single Subject Interventions

Studies reviewed in this section are characterized by only minimal teacher training by the investigators. In these studies, the investigators are responsible for the design of the intervention, and for data collection. Typically, the teacher is instructed how to respond to pupil behavior, but is given only a limited rationale for doing so.

In a study of the effects of teacher attention on the regressed crawling of a nursery school child (Harris et al. 1964), contingent attention was used to decrease the off-feet behavior of a young girl. During baseline observation, the child was on her feet only 10% of the time. When teacher attention was made contingent on on-feet behavior, this figure rose to 60%. After a reversal phase in which on-feet behavior rapidly decreased, teacher attention was again made contingent on on-feet behavior. The frequency of normal walking rose to a level comparable with that of other children in the school. The two teachers received essentially no instruction in principles of behavior modification, and provided only anecdotal descriptions of changes in the child's behavior.

Teacher attention has also been used to modify isolate behavior (Allen et al. 1964), operant crying (Hart et al. 1964), cooperative play (Hart et al. 1968), and digit reversal (Hasazi and Hasazi 1972). In each of these studies teacher attention was made contingent on
the subjects' appropriate behavior. In each case, teacher attention was shown to be a powerful reinforcer capable of bringing about substantial changes in the frequencies of targeted behaviors. Teacher training in these investigations consisted of instruction of how to and when to dispense and withhold attention, and nothing more.

Broden et al. (1970) reported the use of teacher attention to increase the amount of study time of two retarded boys in adjacent classroom desks. The boys' teacher met with Broden for one hour to discuss reinforcement and extinction procedures. After an initial period of baseline observation, the teacher was instructed to attend to the study behavior of only one of the boys, and later, only to the other boy's studying. Following a period of no attention to either boy, the teacher was instructed to attend to the study behavior of both boys. In periods when only one boy was attended to while studying, frequency of study behavior rose greatly for that boy, and to a lesser extent for the boy not attended to. This "spread of effect" is a potentially important by-product of individual interventions, although few investigators have reported it.

Single Subject and Classroom Interventions with Increased Teacher Participation

This second group of studies are reports of interventions in which teachers receive more extensive training and/or take a greater part in actually modifying pupil behavior and collecting data. The focus of studies reviewed in this section is not training per se, but research or remediation of specific behavior problems.
Consultation with teachers regarding specific problems in their classroom management appears to be a popular vehicle for providing teachers with limited training in behavior modification. Perhaps because consultation tends to be a service-related rather than research-oriented phenomenon, however, few well-documented accounts of behavioral consultation with teachers appear in the literature. Of the few relevant accounts available, Fullmer (1972) provides a good example of behavior modification training via consultation. Over a one-week period, Fullmer trained the teacher of eight educable retarded teen-agers to use tokens and contingent praise in the classroom. Pupil behaviors targeted for acceleration were saying words aloud (sight vocabulary), telling time, reading the alphabet and writing numerals. Each correct response was rewarded with praise and tokens which could be redeemed for activities or edibles. At the end of the one-week intervention period, all pupils showed significant gains in the frequencies of target behaviors. Seven months later, the effects of the original training and intervention were being maintained by the teacher's continued use of tokens and contingent praise.

Sulzbacher and Houser (1968) consulted with the teacher of a classroom of retarded children, and taught her how to apply a simple group contingency to eliminate the use of the "naughty finger" in her classroom. After obtaining a baseline frequency of the undesirable gesture, the teacher was instructed to tell her class that each observed occurrence of the behavior would subtract one minute from the time allowed for a special daily recess. Each day while this contingency was in effect, the number of minutes of recess remaining was displayed
to the class on a set of large cards. During the intervention, the frequency of the "naughty finger" dropped from a baseline rate of 17 per day to less than three per day, according to data collected by the teacher. This study demonstrates how a teacher can be given minimal instruction in the use of an isolated behavior modification technique and make major changes in the frequency of a rather common but annoying classroom behavior.

Randolph and Wallin (1973) compared the effects of two consultation procedures on the on-task behavior of nonretarded sixth-graders. After a period of baseline observation of pupils' on-task behavior, teachers were provided with ten weeks of consultation that stressed the use of contingent praise and contingent ignoring (extinction). Throughout the consultation period, teachers were shown data on their own rates of appropriate attending, as well as data reflecting changes in their pupils' on-task behavior. Pupils of half of the teachers participated in out-of-class sessions in which the consultants modeled and reinforced on-task behavior. Both consultation alone and consultation plus modeling and reinforcing on-task behavior led to significant increases in on-task behavior, as compared to a no-treatment condition. No differences were found between the results of the two consultation procedures, suggesting that the controlling contingencies were those managed in class by the teachers, as opposed to those managed out of class by the consultants.

Cooper, Thompson and Baer (1970) investigated the effects of written feedback on the frequency of teachers' attention to appropriate and inappropriate pupil behavior. Feedback, in the form of written
notes was given to the teachers on an hourly basis. Results showed that by providing teachers with periodic feedback about their attending behavior, substantial increases can be made in the frequency of attention to appropriate behavior. In a similar study, Dickerman (1972) trained teachers to code their pupils' behavior as "appropriate" or "inappropriate", and to code their own attending responses as "attending to appropriate behavior" or "attending to inappropriate behavior". Trained observers checked the reliability of the teachers' data throughout the eight-week study. This simple procedure for producing self-generated feedback substantially increased the amount of time students behaved appropriately as it increased the frequency of appropriate teacher attention.

Hall, Lund and Jackson (1968) combined instruction in behavior management, reading assignments, prompting, and daily consultation to train teachers in the use of contingent attention to increase study behavior in the classroom. Results of this investigation are presented as six single subject studies, each with a baseline, contingent attention phase, reversal, and contingent attention phase format. In five of these studies, training was supplemented with in-class cueing of the teacher as to when to attend to pupil behavior. Cueing was not needed in the sixth case. Results show that study behavior was brought under control of teacher attention in all six cases, increasing during periods of contingent attention, and decreasing during reversals. These results are particularly impressive, as the investigation was conducted in a ghetto school known for its children with severe behavior problems such
as fighting, screaming, desk throwing and so on. That simple management
of attention contingencies can bring about significant changes in study
behavior in such a classroom is a forceful demonstration of the rein­
forcing power of teacher attention.

Orme and Purnell (1970) trained two elementary school teachers
of an out-of-control classroom in verbal, nonverbal and token rein­
forcement procedures and in ignoring. In addition, the teachers learned
how to elicit appropriate, reinforceable behavior by asking questions and
by varying stimuli in the classroom.

The teachers also received training in specifying goals of
instruction in terms of pupil behavior, the differential
reinforcement of pupil time on task behavior, pupil silence,
handraising, pupil-pupil cooperation, pupil comments, ques­
tion and answers, pupil attending to another pupil discussing
lesson content, and related educationally relevant behaviors.
In general, the teacher was encouraged to engineer desirable
pupil behaviors that were incompatible with disruptive pupil
behaviors, and then to reinforce the former with tokens and
verbal and nonverbal reinforcement. They were directed to
ignore disruptive behavior (short of fighting, which did not
recur during the experimental period) by focusing on adjacent
pupils modeling desirable behavior (Orme and Purnell 1970, p. 27).

Two classrooms were used in this study. Room A was set up and
conducted as a traditional classroom. Room B was physically restructured
to permit maximum teacher-pupil contact; a token system and a token
store were set up in Room B. Initially, all pupils were exposed to both
classrooms for several days. Later, appropriate behavior in Room A led
to admission to Room B the following day. Results of the study showed
an overall increase of 30% in pupils time-on-task behavior. In Room B
this increase is attributed to the effects of the token system. In
Room A, the increase is attributed partially to generalization effects
but also to the ability of the pupils to delay gratification, and work for a chance to work for reinforcement in Room B the following day.

Hall, Panyan, Rabon and Broden (1968) conducted two experiments in establishing more effective teacher control in a junior high special education class. The teacher was only briefly instructed in principles of behavior modification, but was responsible for carrying out the actual interventions. The goal of both experiments was to increase the amount of classroom study behavior, which occurred 29% of the time during baseline. In the first experiment, the teacher was instructed to verbally and physically praise studying, which rose to 66%. When the teacher began to reward in-seat behavior on a variable interval 18-minute schedule, studying rose to 76%. Finally, a token system was established, and students engaged in study behavior 90% of the time they were supposed to. The second experiment began with a period of baseline observation, during which the token system was discontinued, and during which studying dropped to 40%. When the token system was reinstated, studying rose to 85% and then stabilized at 90%. With the increases in study behavior, the teacher reported decreases in disobedience, cursing, fighting and other disruptive behaviors.

Comprehensive Teacher Training Programs

Next to be reviewed are programs designed to train teachers in a wide range of behavioral principles and techniques. These programs are differentiated from studies reviewed above by the scope of training offered, and by their focus on teachers' skill acquisition, as opposed to immediate remediation of behavior problems.
Hall (1971) has developed a training program that is based on a graduate-level class he conducts for elementary school teachers. Hall notes, however, that a strictly didactic approach to behavior management training is probably not adequate; a program should also include a good deal of practice in actually applying behavior management skills in the classroom. Hall suggests the following guidelines for an effective teacher training program: (1) a classroom lecture series that covers the basics of social learning theory and behavior modification; (2) required reading of basic studies in the behavior modification literature; (3) training in observing and recording behavior; (4) requiring teachers to conduct at least one behavioral intervention during training; (5) providing teachers with opportunities to compare notes on their attempts to modify pupil behavior, so as to increase their exposure to applications of behavior modification in the classroom. Hall reports the results of several interventions that were planned and conducted by his teachers during training. He has found that teachers are capable of conducting their interventions as experiments, and can handle rather rigorous data collection procedures.

Hall reports a typical intervention involving a teen-ager who frequently disputed his teacher's statements and argued with her. Baseline observation by the teacher fixed the rate of disputes at eight per day. When the teacher began ignoring the boy when he tried to argue with her, the rate of disputes dropped to one per day. During a reversal, the teacher again argued with the boy when he disputed her statements, increasing the number of disputes to six per day. In the
final phase of intervention, the teacher again ignored disputes, decreasing them to about one per day. For other examples of teachers' interventions see Hall (1971) and Hall et al. (1971).

Cantrell (1970) investigated the effects of inservice training in operant techniques and videotape feedback on several pupil and teacher behaviors. One experimental group received didactic training plus periodic consultation based on videotapes of classroom interactions. A second group received consultation only, and a third group received neither training nor consultation. As a result of training, both experimental groups of teachers showed significant gains in knowledge of behavior management principles, but no changes in pupil behavior were noted. Lack of change in pupil behavior may be related to the fact that teachers were not required to conduct interventions in the classroom.

In a similar study of inservice training and video feedback, Kosier (1971) instructed each teacher in his group to focus on the two children in her classroom who presented the most behavior problems. Pre- and posttraining measures were taken on the frequency of teachers' verbal praise and reprimands; percent on-task time for all pupils in each class was recorded throughout the investigation. While the frequency of teachers' verbal praise increased over the training period, the amount of on-task behavior did not increase for the group of pupils as a whole, which is similar to Cantrell's (1970) findings. For the identified problem children, however, percent on-task time increased from a baseline of 58% to a high of 78% in the group whose teachers received training only, and from a similar baseline to 88% in the pupils of teachers who received training plus video feedback. In
comparing these results with those of Cantrell (1970), it can be spec-
ulated that the crucial difference between the two is that Kosier's
procedures focused the teachers' attention on a small number of children,
and thus probably focused the effects of training. This suggests that
training should initially emphasize "thinking small" and focusing on
a limited number of children and behaviors, before asking teachers to
conduct interventions involving many children and/or many behaviors.

Zimmerman and Zimmerman (1971) suggest that it is important to
present a few basic skills to teachers in a lecture format, but that
each teacher be allowed to decide which techniques of behavior modifica-
tion she wants to use in her classroom. That is, a teacher should be
encouraged to "do her own thing" and the trainers should offer sugges-
tions for doing her own thing with precision. For example, if a teacher
thinks that a token system is akin to bribing a child for being good,
she should not necessarily be dissuaded from this opinion. Rather,
the trainer should offer the teacher an alternate technique that is
not objectionable to her, e.g., social reinforcement. The primary goals
of the Zimmermans' program are: (1) learning to define problems ob-
jectively; (2) applying self-chosen behavior modification techniques
to academic and conduct problems; (3) learning to quantitatively assess
outcomes.

Thirteen teachers participated in the Zimmermans' program.
Nine of the teachers elected to receive college credit for attending,
and four elected to receive fifty dollars. It is interesting to note
that each of the teachers receiving credit completed at least two inter-
ventions, while those receiving money attempted only one intervention.
apiece—a finding that might be useful to future trainers with university connections. One year after the completion of training, ten of the thirteen teachers were still teaching school. Nine indicated that they had used their training at least once during the previous year and seven of the nine were engaged in behavior management projects at the time of follow-up.

The final training program to be considered is the Behavior Management Program for Teachers (Stone et al. 1973). This manual is the basis for a previously unresearched training program that has been conducted by the staff of the Tucson Child Guidance Center for training teachers in the Tucson, Arizona public schools. Based on an earlier manual for training parents (Stone, Hanson and Evans 1972), the teacher training manual is a sequential guide to acquisition of classroom behavior management skills. Along with sections covering social learning theory and behavioral intervention, the manual includes recommendations for in-class projects that give teachers practice in applying behavior management skills. The material in the manual is organized so that teachers first learn groundwork skills such as observing and recording; more sophisticated behavior modification skills such as shaping are presented as training progresses. The manual is written to be used in conjunction with didactic training by behaviorally-oriented mental health professionals, and was used in the present study.
CHAPTER 3

PURPOSE AND CONTEXT OF THE PRESENT STUDY

The present study was planned and designed while the author was a clinical psychology extern at the Tucson Child Guidance Center in Tucson, Arizona. Prior to the initiation of the study, the Center's teacher training program had been offered in six elementary schools in the Tucson area. Approximately sixty teachers had participated in the program, all of whom taught elementary classrooms of nonretarded children. Anecdotal reports from, and data collected by teachers who had participated in the program indicated that it is an effective means for training teachers to modify the behavior of their pupils. This is consistent with findings of investigators who have conducted similar training programs, e.g., Hall (1971), Kosier (1971), and Zimmerman and Zimmerman (1971).

While the Behavior Management Program for Teachers appeared to be effective, it had not been researched in any methodical manner. The first purpose of the present study was to assess the effects of the training program on program participants and their pupils. More specifically, the present study was designed to compare the effects of purely didactic training to the effects of didactic training plus in-class practice of behavior management skills. The training manual is structured so that during the first half of training, which takes four weeks, teachers attend training sessions and collect data in the
classroom, but do not attempt to modify pupil behavior. During the second four weeks of training, teachers continue to attend training sessions and collect data and also attempt several interventions in the classroom, using the behavior management skills they have learned. By comparing changes in teacher and pupil behavior during the first half of training to changes during the second half, inferences can be drawn concerning the relative effects of didactic training and in-class practice of behavior management skills.

A second purpose of the present study was to discover whether or not the program is effective in training teachers of retarded pupils. To date there are no published accounts of comprehensive behavior modification training programs for teachers of the retarded. Thus, the present study is an investigation of the effects of didactic training and in-class practice of behavior management skills on teachers of the retarded and their pupils. The program described herein was conducted by two social workers and two psychologists, and was offered to the school in which the research was conducted as a free service of the Tucson Child Guidance Center and the Tucson East Mental Health Center.
Subjects were randomly selected from a group of 21 teachers who expressed an interest in attending the training program. Eight subjects were selected, all of whom agreed to attend all training sessions and all of whom agreed to permit observers in their classrooms. The remaining 13 teachers attended the training sessions but did not participate in the research. Subjects were all adaptive education teachers at Canoung School in Tucson, Arizona. Subjects' range of teaching experience was from 1.5 years to 8.0 years with a mean of 4.0 years.

Pupils of the subjects were all previously diagnosed as "trainable mentally retarded" (I.Q. range from 25 to 50). Exact I.Q. scores could not be obtained for reasons of confidentiality. All of the pupils were ambulatory and had adequate self-help skills. The eight classes were divided into two groups on the basis of pupils' age, four for older children and four for younger ones. In actuality, the age ranges of the two groups overlapped considerably, with the age range for the "younger" classes being from 5.5 years to 9.5 years and the range for the "older" classes being from 6.5 years to 13.5 years. Mean ages of pupils in the two sets of classes were 7.3 years and 7.7 years respectively; the overall mean age was 7.5 years.
General Procedures

Subjects met with the trainers two weeks prior to the beginning of training to discuss the mechanics of the project, to arrange weekly observation times, and to complete a brief questionnaire covering basic principles of behavior modification. Baseline data collection began during the week of this first meeting. For each week thereafter, observers collected observational data in each classroom for 30 minutes per week. After a two-week baseline period, teachers met with the trainers once a week for approximately 90 minutes. Training lasted for eight weeks. Following the final week of training, teachers met once again with the trainers to discuss the program, to receive preliminary summaries of data collected, and to again complete the brief questionnaire.

Each training session began with a half-hour lecture that closely paralleled the material in the chapter of the manual to be covered that week. Chapters of the manual were handed out at the end of each session. During the lecture portion of the session, trainers often performed "behavioral demonstrations" (skits) which portrayed techniques of classroom behavior management. When subjects asked questions during the lecture, they were answered briefly if germane to the topic being presented. Questions not relating to the lecture material were deferred until after the lecture, or answered with, "That's a good question, and we'll talk about it when we get to the chapter that covers that topic." Every effort was made to follow the format and order of the manual as closely as possible. The topics covered in lecture each week are as follows:
Week 1  Overview of behavior modification, specifying behavioral objectives;
Week 2  Techniques to increase appropriate behavior;
Week 3  Using reinforcers to change behavior;
Week 4  Techniques to decrease problem behavior;
Week 5  Shaping, modeling;
Week 6  Rules, peer reinforcement;
Week 7  Contingency management plans for the whole class;
Week 8  Punishment, problems at home, problem-solving, arming oneself against skeptics.

A more complete description of weekly topics and training procedures can be found in Appendix B. Following the lecture and a ten-minute discussion period, individual trainers met with groups of two or three teachers to inspect data collected the previous week, to discuss upcoming assignments, and to answer questions about specific behavior management problems relating to the lecture material.

**Data Collection**

During each observation session, the observer counted the frequency of and coded teacher attention and antecedent pupil behavior. Teacher attention was coded as Verbal-neutral (V), Verbal Praise (V+), Verbal Reprimand (V-), Physical-neutral (P), Physical Praise (P+), or Physical Punishment (P-). Antecedent pupil behavior was coded as Appropriate Verbal (AV), Inappropriate Verbal (IV), Appropriate Physical (AP), or Inappropriate Physical (IP). In cases where teacher attention appeared to be elicited by more than one pupil, the antecedent pupil behavior was recorded as "Group" (G).
Each observation session was divided into 6 five-minute blocks. During three of the blocks the observer coded interactions between the teacher and all members of the class, excluding one child the teacher had identified to the investigators as her biggest behavior-problem child. During the other three blocks of time the observer coded interactions between the teacher and this designated behavior-problem child. Definitions of the coding categories are listed in Appendix A.

Three observers collected the observational data. All three had used the recording system during the piloting of the system in several special education classrooms in Tucson. During the pilot study and the present research, inter-observer agreement was calculated as percent agreement (the number of agreements in a five-minute block divided by number of agreements plus number of disagreements). During the pilot study, percent agreement was 85% for observers 1 and 2, 89% for observers 1 and 3, and 87% for observers 2 and 3.

Throughout the present study, observers 1 and 2 were assigned to collect data in four classrooms each. During the second, fifth and ninth weeks of the study, however, observers 1 and 2 made simultaneous observations in four classrooms. During weeks three and seven, observer 3 made reliability checks in all eight classrooms. This rather complex schedule of reliability checks was predicated solely on the schedule requirements of the three available observers. Overall reliability was 82%. Table 1 lists reliability figures for each week checks were made; each entry represents percent agreement between two observers, averaged over four classrooms.
Table 1. Reliability of Observational Measures

<table>
<thead>
<tr>
<th>Observers</th>
<th>Week</th>
<th></th>
<th></th>
<th></th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 and 2</td>
<td>2</td>
<td>.83</td>
<td></td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>--</td>
<td></td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>1 and 3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 and 3</td>
<td>2</td>
<td></td>
<td></td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2</td>
<td>.83</td>
<td></td>
<td></td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.815</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td>.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>.78</td>
<td></td>
</tr>
</tbody>
</table>
Prior to the first training session and immediately following
the final meeting with the trainers, each subject completed a 27-item
multiple choice questionnaire covering principles and techniques of
behavior modification presented during training. The questionnaire
is reproduced in Appendix C.

Beginning with the second week of training, teachers were as­
signed tasks that required them to collect data in class each day.
Teachers were encouraged to keep collecting data as they tried suggested
behavior modification techniques during the course of training.

Data Analysis

The observational data were analyzed as twenty separate de­
pendent measures (ten coding categories times two observational targets,
teachers and targeted behavior-problem children, and teachers and
non-targeted children). For each measure, weekly frequencies were an­
alyzed as a one-factor ANOVA with ten repeated measurements.

A problem inherent in repeated measurements ANOVA's is the
confounding of possible Subject X Trials interaction effects with the
error term. The existence of significant Subject X Trials interactions
in a set of data is referred to as "nonadditivity" (Meyers 1966) and
can result in an invalid F ratio for the Trials effect. Meyers suggests
that any set of data to be analyzed as a repeated measurements ANOVA
first be tested for nonadditivity. If the nonadditivity is significant,
Meyers suggests that the transformation $\log_{10}(X+1)$ will often correct
for nonadditive effects.
Following Meyers' recommended strategy, the set of data for each of the twenty observational measures was analyzed in the following manner. First, the ANOVA was performed on the original data. Second, if the $F$ ratio for Trials was significant, the data was tested for nonadditivity to determine whether or not the $F$ ratio could be considered valid. Third, the data was transformed using the $\log_{10}(X+1)$ transformation, and the ANOVA was repeated. Fourth, if the $F$ ratio for Trials using the transformed data was significant, the transformed data was tested for nonadditivity. If transformation had indeed corrected for nonadditivity and tests of nonadditivity, the 0.05 level was accepted (two-tailed test).

The questionnaire data was also analyzed for ANOVA with two repeated measurements. These were analyzed for nonadditive effects.

**Hypotheses**

This study is designed to test 21 specific hypotheses concerned with changes in the behavior of the subjects (teachers) and their pupils. Each hypothesis is listed in its null form.

**Hypothesis 1.** Mean weekly frequencies of teachers' Verbal Praise ($V+$) of the targeted behavior-problem children will not differ significantly.

**Hypothesis 2.** Mean weekly rates of teachers' Verbal Reprimands ($V-$) of targeted children will not differ significantly.
Hypothesis 3. Mean weekly rates of teachers' Verbal-neutral (V) statements to targeted children will not differ significantly.

Hypothesis 4. Mean weekly rates of teachers' Hypothesized Positive (P+) of targeted children will not differ significantly.

Hypothesis 5. Mean weekly rates of teachers' Hypothesized Negative (P-) of targeted children will not differ significantly.

Hypothesis 6. Mean weekly rates of teachers' Hypothesized Positive (P) behaviors directed toward targeted children will not differ significantly.

Hypothesis 7. Mean weekly rates of targeted children's Appropriate Verbal (AV) behavior will not differ significantly.

Hypothesis 8. Mean weekly rates of targeted children's Appropriate Physical (AP) behavior will not differ significantly.

Hypothesis 9. Mean weekly rates of targeted children's Inappropriate Physical (IP) behavior will not differ significantly.

Hypothesis 10. Mean weekly rates of targeted children's Inappropriate Verbal (IV) behavior will not differ significantly.

Hypotheses 11-20 are identical to hypotheses 1-10, but pertain to interactions between teachers and non-targeted pupils.

The final hypothesis is concerned with the questionnaire scores. Hypothesis 21. Pre- and posttreatment means of questionnaire scores will not differ significantly.
Following Meyers' recommended strategy, the set of data for each of the twenty observational measures was analyzed in the following manner. First, the ANOVA was performed on the original data. Second, if the $F$ ratio for Trials was significant, the data was tested for nonadditivity to determine whether or not the $F$ ratio could be considered valid. Third, the data was transformed using the $\log_{10}(X+1)$ transformation, and the ANOVA was repeated. Fourth, if the $F$ ratio for Trials using the transformed data was significant, the transformed data was tested for nonadditivity to see if the transformation had indeed corrected for nonadditive effects. For all ANOVA's and tests of nonadditivity, the 0.05 level of significance was used (two-tailed test).

The questionnaire data was analyzed as a one-factor ANOVA with two repeated measurements. This data was also analyzed for nonadditive effects.

Hypotheses

This study is designed to test 21 specific hypotheses concerned with changes in the behavior of the subjects (teachers) and their pupils. Each hypothesis is listed in its null form.

Hypothesis 1. Mean weekly frequencies of teachers' Verbal Praise ($V^+$) of the targeted behavior-problem children will not differ significantly.

Hypothesis 2. Mean weekly rates of teachers' Verbal Reprimands ($V^-$) of targeted children will not differ significantly.
Hypothesis 3. Mean weekly rates of teachers' Verbal-neutral (V) statements to targeted children will not differ significantly.

Hypothesis 4. Mean weekly rates of teachers' Physical Praise (P+) of targeted children will not differ significantly.

Hypothesis 5. Mean weekly rates of teachers' Physical Punishment (P-) of targeted children will not differ significantly.

Hypothesis 6. Mean weekly rates of teachers' Physical-neutral (P) behaviors directed toward targeted children will not differ significantly.

Hypothesis 7. Mean weekly rates of targeted children's Appropriate Verbal (AV) behavior will not differ significantly.

Hypothesis 8. Mean weekly rates of targeted children's Appropriate Physical (AP) behavior will not differ significantly.

Hypothesis 9. Mean weekly rates of targeted children's Inappropriate Physical (IP) behavior will not differ significantly.

Hypothesis 10. Mean weekly rates of targeted children's Inappropriate Verbal (IV) behavior will not differ significantly.

Hypotheses 11–20 are identical to hypotheses 1–10, but pertain to interactions between teachers and non-targeted pupils.

The final hypothesis is concerned with the questionnaire scores.

Hypothesis 21. Pre- and posttreatment means of questionnaire scores will not differ significantly.
CHAPTER 5

RESULTS

Results of the analysis of the teacher-pupil observational data (see Data Analysis under METHOD, CHAPTER 4) are summarized in Table 2. Of six significant treatment effects found in the original data, five are possibly invalid due to significant nonadditive effects. For the transformed data, only two are possibly invalidated by non-additive effects. As the transformed data appear to fit the "additive" model of the repeated measures ANOVA better, results presented below are for the transformed data.

Null hypothesis 1 is rejected. Mean weekly rates of teachers' Verbal Praise (VP) of targeted pupils differ significantly ($F = 2.45, p < 0.05$). Post-hoc Scheffe tests comparing baseline to weeks 3-6 (didactic training), and baseline to weeks 4-8 (didactic training plus in-class practice) indicate that the significant Trials effect is not due to differences in training. The critical value of a Mean Square in the Scheffe test is 0.8765 at the recommended 0.1 level. The obtained values were 0.0388 and 0.5343. The results of this ANOVA and post-hoc tests appear in Table 3. Weekly mean frequencies of VP of targeted children are plotted in Figure 1.

Null hypothesis 3 is rejected. Mean weekly rates of teachers' Verbal-neutral (V) statements to targeted children differ significantly ($F = 2.205, p < 0.05$). Post-hoc tests revealed no significant effects.
Table 2. Summary of Analysis of Observational Data

N = Non-target children
T = Target children
Significant F ratios are underscored (0.05 level)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Original Data</th>
<th>Transformed Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F_trials</td>
<td>F_nonadditive</td>
</tr>
<tr>
<td>VN</td>
<td>3.328</td>
<td>18.221</td>
</tr>
<tr>
<td>V+N</td>
<td>0.867</td>
<td>-----</td>
</tr>
<tr>
<td>V-N</td>
<td>1.535</td>
<td>-----</td>
</tr>
<tr>
<td>PN</td>
<td>1.771</td>
<td>-----</td>
</tr>
<tr>
<td>P+N</td>
<td>2.399</td>
<td>15.610</td>
</tr>
<tr>
<td>P-N</td>
<td>1.605</td>
<td>-----</td>
</tr>
<tr>
<td>AVN</td>
<td>3.528</td>
<td>16.254</td>
</tr>
<tr>
<td>APN</td>
<td>0.618</td>
<td>-----</td>
</tr>
<tr>
<td>IPN</td>
<td>2.783</td>
<td>24.403</td>
</tr>
<tr>
<td>IVN</td>
<td>0.481</td>
<td>-----</td>
</tr>
<tr>
<td>VT</td>
<td>1.295</td>
<td>-----</td>
</tr>
<tr>
<td>V+T</td>
<td>2.348</td>
<td>4.892</td>
</tr>
<tr>
<td>V-T</td>
<td>0.955</td>
<td>-----</td>
</tr>
<tr>
<td>PT</td>
<td>0.829</td>
<td>-----</td>
</tr>
<tr>
<td>P+T</td>
<td>1.624</td>
<td>-----</td>
</tr>
<tr>
<td>P-T</td>
<td>0.427</td>
<td>-----</td>
</tr>
<tr>
<td>AVT</td>
<td>6.779</td>
<td>0.443</td>
</tr>
<tr>
<td>APT</td>
<td>0.943</td>
<td>-----</td>
</tr>
<tr>
<td>IPT</td>
<td>1.416</td>
<td>-----</td>
</tr>
<tr>
<td>IVT</td>
<td>1.316</td>
<td>-----</td>
</tr>
</tbody>
</table>
Table 3. Summary of ANOVA, Verbal Praise of Targeted Children

B = Baseline  
E1 = First half of training  
E2 = Second half of training  
MS<sub>crit</sub> = Critical value of a mean square (0.1 level)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>2.701</td>
<td>7</td>
<td>0.296</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>1.277</td>
<td>9</td>
<td>0.136</td>
<td>2.454</td>
<td>0.05</td>
</tr>
<tr>
<td>Error</td>
<td>3.505</td>
<td>63</td>
<td>0.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6.853</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post hoc

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS&lt;sub&gt;crit&lt;/sub&gt;</td>
<td>---</td>
<td>9</td>
<td>0.8769</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS&lt;sub&gt;B vs. E1&lt;/sub&gt;</td>
<td>---</td>
<td>9</td>
<td>0.3888</td>
<td>---</td>
<td>NS</td>
</tr>
<tr>
<td>MS&lt;sub&gt;B vs. E2&lt;/sub&gt;</td>
<td>---</td>
<td>9</td>
<td>0.5343</td>
<td>---</td>
<td>NS</td>
</tr>
</tbody>
</table>
Fig. 1. Teachers' Verbal Praise of Targeted Children
for type of training. The summary of this ANOVA and results of post-hoc tests appear in Table 4. Mean weekly frequencies of V statements to targeted children are plotted in Figure 2.

Null hypothesis 7 is rejected. Mean weekly frequencies of targeted children's Appropriate Verbal (AV) behavior differ significantly ($F = 10.637, p < 0.001$). Post-hoc tests reveal significant effects for didactic training versus baseline, and for didactic training plus in-class practice versus didactic training alone. Results of the ANOVA appear in Table 5, and are plotted in Figure 3.

Null hypothesis 17 is rejected. Mean weekly frequencies of non-targeted children's Appropriate Verbal (AV) behavior differ significantly ($F = 2.232, p < 0.05$). Post-hoc tests reveal no significant effects for type of training. The summary table of the ANOVA appears in Table 6. Weekly frequencies are plotted in Figure 4.

Null hypothesis 19 is rejected. Mean weekly frequencies of non-targeted children's Inappropriate Physical (IP) behavior differ significantly ($F = 2.622, p < 0.05$). Post-hoc tests reveal no significant effects for type of training. The summary of the ANOVA appears in Table 7, and the weekly frequencies are plotted in Figure 5.

Null hypotheses 2, 5, 6, 8, 9, 10, 11, 12, 13, 15, 16, 18, and 20 are retained. None of these measures showed significant differences in their weekly frequencies.

Tests of null hypotheses 4 and 14 are not considered valid, due to significant nonadditive effects in the transformed data.
Table 4. Summary of ANOVA, Verbal-neutral Statements to Targeted Children

B = Baseline  
E1 = First half of training  
E2 = Second half of training  
$MS_{\text{crit}}$ = Critical value of a mean square (0.1 level)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>0.322</td>
<td>7</td>
<td>0.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>1.544</td>
<td>9</td>
<td>0.172</td>
<td>2.205</td>
<td>0.05</td>
</tr>
<tr>
<td>Error</td>
<td>4.902</td>
<td>63</td>
<td>0.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6.768</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post hoc

- $MS_{\text{crit}}$  
- $MS_B$ vs. E1  
- $MS_B$ vs. E2
Fig. 2. Teachers' Verbal-neutral Statements to Targeted Children
Table 5. Summary of ANOVA, Appropriate Verbal Behavior of Targeted Children

B = Baseline
E1 = First half of training
E2 = Second half of training
$MS_{crit}$ = Critical value of a mean square (0.1 level)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>2.286</td>
<td>7</td>
<td>0.327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>4.988</td>
<td>9</td>
<td>0.544</td>
<td>10.637</td>
<td>0.001</td>
</tr>
<tr>
<td>Error</td>
<td>3.282</td>
<td>63</td>
<td>0.052</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10.556</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post hoc

$MS_{crit}$

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$MS_{B \ vs. \ E1}$</td>
<td>---</td>
<td>9</td>
<td>0.9688</td>
<td>---</td>
<td>0.05</td>
</tr>
<tr>
<td>$MS_{B \ vs. \ E2}$</td>
<td>---</td>
<td>9</td>
<td>4.4450</td>
<td>---</td>
<td>0.01</td>
</tr>
<tr>
<td>$MS_{E1 \ vs. \ E2}$</td>
<td>---</td>
<td>9</td>
<td>1.8900</td>
<td>---</td>
<td>0.01</td>
</tr>
<tr>
<td>$MS_{linear}$</td>
<td>---</td>
<td>9</td>
<td>1.9273</td>
<td>---</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Fig. 3. Appropriate Verbal Behavior of Targeted Children
Table 6. Summary of ANOVA, Appropriate Verbal Behavior of Non-targeted Children

B = Baseline
E1 = First half of training
E2 = Second half of training
MS crit = Critical value of a mean square (0.1 level)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>5.560</td>
<td>7</td>
<td>0.795</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>1.530</td>
<td>9</td>
<td>0.170</td>
<td>2.232</td>
<td>0.05</td>
</tr>
<tr>
<td>Error</td>
<td>4.790</td>
<td>63</td>
<td>0.760</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11.880</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post hoc

| MS crit         | --- | 9  | 1.1901|      |       |
| MS B vs. E1     | --- | 9  | 0.2952| ---  | NS    |
| MS B vs. E2     | --- | 9  | 0.3247| ---  | NS    |
Fig. 4. Appropriate Verbal Behavior of Non-targeted Children
Table 7. Summary of ANOVA, Inappropriate Physical Behavior of Non-targeted Children

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>1.691</td>
<td>7</td>
<td>0.242</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>2.113</td>
<td>9</td>
<td>0.234</td>
<td>2.622</td>
<td>0.05</td>
</tr>
<tr>
<td>Error</td>
<td>5.630</td>
<td>63</td>
<td>0.089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9.424</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post hoc

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{MS}_{\text{crit}}$</td>
<td>---</td>
<td>9</td>
<td>1.3937</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{MS}_B$ vs. $\text{E}1$</td>
<td>---</td>
<td>9</td>
<td>0.2306</td>
<td>---</td>
<td>NS</td>
</tr>
<tr>
<td>$\text{MS}_B$ vs. $\text{E}2$</td>
<td>---</td>
<td>9</td>
<td>1.2470</td>
<td>---</td>
<td>NS</td>
</tr>
<tr>
<td>$\text{MS}_{\text{linear}}$</td>
<td>---</td>
<td>9</td>
<td>1.7430</td>
<td>---</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Fig. 5. Inappropriate Verbal Behavior of Non-Targeted Children
Null hypothesis 21 is rejected. Posttest scores on the questionnaire were significantly higher than the pretest scores ($F = 13.194, p < 0.01$). See Table 8.

Subjects began taking data on their pupils' behavior during the second week of training. Throughout the second half of training, weekly assignments included instructions for the subjects to attempt to modify at least one behavior of one pupil. While all eight subjects verbally reported that they were conducting an intervention in the classroom, and while each subject presented some data during at least one of the training sessions, only three subjects actually turned in final reports of projects attempted. Each project is discussed briefly below.

Figure 6 shows the frequency of eight-year-old David's use of the "naughty finger". (The "naughty finger" is a deprecatory gesture that consists of extension of the middle finger and flexion of the other fingers downward toward the palm.) David's teacher reported that she and the pupils in the class tended to react strongly to this gesture and that David received a lot of attention for using it. The teacher collected baseline data that showed the frequency of the gesture to be about eight per day. When the teacher began ignoring David's use of the gesture, the frequency dropped to slightly less than seven per day. Starting on the sixteenth day of the intervention, the teacher instructed the class to ignore David when he used the gesture. The frequency of the "naughty finger" dropped to about 5.5 per day by the end of the training program. Evidently, the gesture was being maintained by attention from pupils and the teacher.
Table 8. Summary of ANOVA, Questionnaire Scores

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>58.00</td>
<td>8</td>
<td>7.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>76.00</td>
<td>1</td>
<td>76.00</td>
<td>13.14</td>
<td>0.05</td>
</tr>
<tr>
<td>Error</td>
<td>45.00</td>
<td>8</td>
<td>5.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>179.00</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 6. David's Use of the "Naughty Finger"
The next two projects to be reported were conducted by one teacher, and involve an eight-year-old girl named Vera. The teacher complained that Vera would not follow simple directions such as "Put on your coat," or "Get your crayons," even though she was capable of doing so. During baseline, Vera was following only about three directions per day, or about half of the directions given her. During the intervention phase, the teacher contingently praised Vera for following directions. The rate of following directions rose steadily, as did the teacher's rate of giving directions. By the end of the training program, Vera was following 90% of all directions at a rate of nine per day. (See Figure 7.)

Vera also had a habit of twisting her nose and putting her hands in her mouth. During baseline, these two behaviors occurred at a rate of 22 per day. The first intervention tried by the teacher was a procedure she called "restitution", which consisted of requiring Vera to wash her hands and/or wipe her nose with a tissue each time she twisted her nose or put her hand in her mouth. During this phase, the frequency of the two behaviors dropped to 15 per day. Finally, the teacher instituted a procedure called "Full Motor Training" (FMT), which she made contingent on nose-twisting and hand-mouthing. FMT consisted of approximately two minutes of exercises involving the arms, hands, and head. By the end of the training program, the frequency of nose-twisting and hand-mouthing was approximately 5 per day. Figure 8 shows the results of this intervention.

The final project to be reported involves handraising. The teacher reported that his eight pupils did not raise their hands before
Fig. 7. Vera: Following Directions
Fig. 8. Vera: Twisting of Nose and Hands in Mouth
asking questions during writing period. The teacher collected data on handraising, and fixed the baseline rate at two per thirty-minute period. To try to increase the frequency of handraising, he decided to ignore children who asked questions without first raising their hands. The frequency of handraising rose sharply to a high of nine times per period by the end of the program. Although the teacher did not keep data on the frequency of talking out, he indicated that this behavior decreased as handraising increased, to the point where talking without first raising a hand had been virtually eliminated. The daily frequencies of handraising are presented in Figure 9.

Time and manpower limitations did not permit classroom observers to check the reliability of teachers' projects data. The observers did, however, make brief notes about the projects reported above. On David's use of the "naughty finger", an observer noted, "There seems to be a slight decrease in this behavior, but the teacher is recording only about one-third of its occurrences". About Vera, an observer noted that FMT had greatly reduced the rate of nose twists. In regard to the handraising project, an observer noted that ignoring talking-out had virtually eliminated this behavior, and had increased handraising to near 100%.

As a measure of relationship between observational coding categories, frequencies of observations in successive five-minute blocks for each category were correlated. Two matrices of Pearson's product-moment correlations were generated, one for data on interactions between teachers and targeted pupils, and one for data on interactions between teachers and non-targeted pupils. These two
Fig. 9. Handraises during Writing Period
matrices appear in Table 9 and Table 10 respectively. Critical values for determining significance of correlations were interpolated from a table of such values (Guilford 1956, p. 539).

For interactions between teachers and targeted behavior-problem pupils, pairs of observational measures that correlate significantly and positively are V and AP, V and AV, V+ and AV, V+ and P, P and P+, P+ and AP, AV and AP. For interactions between teachers and the other non-targeted pupils, significant positive correlations were found for the pairs V and AP, V and AV, V and V+, V+ and AP, AV and AP. Significant negative correlations were found for the pairs V and IP, V+ and IP, P+ and IP, P+ and P−, IP and AV, IP and AP. The 95% confidence level was used in judging the significance of correlations.
Table 9. Correlations between Observational Measures, Behaviors of Teachers and Targeted Children

Significant correlations (0.05 level) are underscored

<table>
<thead>
<tr>
<th>Measure</th>
<th>IV</th>
<th>IP</th>
<th>AP</th>
<th>AV</th>
<th>P-</th>
<th>P+</th>
<th>P</th>
<th>V-</th>
<th>V+</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>0.014</td>
<td>-0.373</td>
<td>0.668</td>
<td>0.864</td>
<td>-0.419</td>
<td>0.326</td>
<td>0.514</td>
<td>-0.390</td>
<td>0.449</td>
<td>---</td>
</tr>
<tr>
<td>V+</td>
<td>-0.63</td>
<td>-0.290</td>
<td>0.517</td>
<td>0.709</td>
<td>0.125</td>
<td>0.363</td>
<td>0.709</td>
<td>-0.238</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>V-</td>
<td>-0.428</td>
<td>0.358</td>
<td>-0.105</td>
<td>-0.261</td>
<td>-0.386</td>
<td>0.107</td>
<td>-0.351</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>0.088</td>
<td>0.056</td>
<td>0.603</td>
<td>0.572</td>
<td>-0.088</td>
<td>0.690</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P+</td>
<td>-0.469</td>
<td>-0.053</td>
<td>0.747</td>
<td>0.538</td>
<td>-0.262</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-</td>
<td>0.093</td>
<td>0.160</td>
<td>-0.466</td>
<td>-0.267</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AV</td>
<td>-0.280</td>
<td>-0.330</td>
<td>0.857</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP</td>
<td>-0.380</td>
<td>-0.250</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IP</td>
<td>-0.048</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
Table 10. Correlations between Observational Measures, Behaviors of Teachers and Non-targeted Children

Significant correlations (0.05 level) are underscored

<table>
<thead>
<tr>
<th>Measure</th>
<th>IV</th>
<th>IP</th>
<th>AP</th>
<th>AV</th>
<th>P-</th>
<th>P+</th>
<th>P</th>
<th>V-</th>
<th>V+</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>-0.019</td>
<td>-0.759</td>
<td>0.919</td>
<td>0.837</td>
<td>-0.219</td>
<td>0.547</td>
<td>0.229</td>
<td>-0.270</td>
<td>0.650</td>
<td>---</td>
</tr>
<tr>
<td>V+</td>
<td>0.020</td>
<td>-0.686</td>
<td>0.739</td>
<td>0.386</td>
<td>-0.001</td>
<td>0.477</td>
<td>0.552</td>
<td>-0.592</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>V-</td>
<td>-0.436</td>
<td>0.316</td>
<td>-0.235</td>
<td>-0.188</td>
<td>-0.186</td>
<td>-0.228</td>
<td>-0.525</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>P</td>
<td>-0.212</td>
<td>-0.247</td>
<td>0.307</td>
<td>-0.176</td>
<td>-0.010</td>
<td>0.149</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>P+</td>
<td>0.408</td>
<td>-0.884</td>
<td>0.409</td>
<td>0.635</td>
<td>-0.745</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>P-</td>
<td>-0.234</td>
<td>0.625</td>
<td>-0.052</td>
<td>-0.512</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>AV</td>
<td>0.173</td>
<td>-0.713</td>
<td>0.668</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>AP</td>
<td>-0.241</td>
<td>-0.649</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>IP</td>
<td>-0.362</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>IV</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
Two methodological issues bear some discussion in relation to the interpretation of the results of the present study, namely, the absence of control groups, and the transformation of the observational data.

An obviously desirable change in the design of the study would have been the inclusion of another treatment group for which types of training were reversed in order, and a no-treatment control group. A second training group would have permitted conclusions to be drawn about effects due to order of training procedures; a no-treatment control group would have controlled for possible "placebo" effects. It can only be said that there were not enough qualified teachers available to serve as control subjects, and not enough trainers to conduct two more groups had subjects been available. While the exclusion of control groups limits the interpretation of results obtained, the literature provides some clues as to what one might expect had two more groups been used.

There are no available investigations of the effects of varying order of didactic training and didactic training plus in-class practice of behavior modification on teachers' skill acquisition. Eyberg (1973) has found, however, that the relative difficulty of assigned interventions does not effect the outcome of parents' attempts
to modify their children's behavior. If the two types of training in the present study can be construed as two levels of difficulty, one might predict that varying their order would have no effect on teachers' acquisition of behavior modification skills. It can also be noted that various investigators (e.g., Sulzbacher and Houser 1968) have found that teachers are capable of modifying pupil behavior at the same time they are receiving instruction in behavior modification techniques. Again, one might predict that reversing the order of training in the present study would have no significant effect. Such prediction by analogy is, of course, only speculative, and does not address the possibility that teachers might actually become better behavior modifiers if initially exposed to didactic training plus in-class practice of behavior modification skills. (The structuring of the present program is based on the assumption that training would be less effective if purely didactic training did not precede in-class practice.) Realizing the value of analogy and assumption, it would seem prudent to limit inferences based on the results of this study in accordance with the training sequence actually used, as opposed to one that might have been used.

In regard to changes in classroom behavior one might expect to observe in a no-treatment control group, the literature provides some more clearly cut evidence. Studies that have compared behavior management training to no training have consistently found a lack of significant change in the behavior of teachers and pupils serving as subjects in no-treatment control groups (Dickerman 1972, Walker and Buckley 1972,
Randolph and Wallin 1973). Thus, there is modest reason to conjecture that had a no-treatment control group been used in the present study, that no significant changes would have been observed in the behavior of the control teachers or their pupils. Again in the name of scientific conservatism, however, specific results of the present study will be discussed with the full realization that the hypothesized lack of effects in the non-included control group cannot be adequately predicted, but only inferred.

The second methodological issue to be considered is the transformation of the observational data. While the $\log_{10}(X+1)$ transformation is appropriate in the statistical sense, it results in a set of data that does not reflect the dependent variables of interest (pupil and teacher behavior) with 100% accuracy. Meyers (1966) notes that most psychologists find transformed data confusing at best and meaningless at worst, unless the relationship between the original and transformed data is known. By referring to Figures 1, 2, 3, and 4, the relationship between the original and transformed data becomes clear. (In each figure the mean weekly frequencies of the original data have been divided by a constant so that they can be presented on the same scales as the transformed data.) As these figures show, both sets of data have similar trends over time for each observational measure. What the figures do not show is the decreased variability within each weekly mean of the transformed data, which is the result of and the rationale for the $\log_{10}(X+1)$ transformation.
One might argue that the utility of the transformed data is severely limited, as the dependent variables of interest are still frequencies of teacher and pupil behaviors and not the frequencies of logarithms. Lindsley (1973) notes that this is a common bit of data-bound nearsightedness in many psychologists, and counters that in any sort of therapeutic endeavor, change is the crucial variable of interest, independent of the scale by which it is measured. As long as data reflects change (or in Lindsley's jargon "celeration") in behavior, all else is secondary. Interestingly, Lindsley advocates the use of logarithmic chart paper for plotting data, as he thinks this graphic transformation better represents gradual changes in behavior than conventional frequency graphs. The present author concurs with Lindsley that change is the most important issue when attempting to modify behavior, and thinks that the transformed data in this study accurately reflect changes in the behavior of the subjects and their pupils.

Of the twenty observational measures of teacher and pupil behavior, five proved to be significantly nonadditive and showed significant changes over the course of training (AV targeted children, IP targeted children, V non-targeted children, V+ non-targeted children, and AV non-targeted children). Of these five, three were significant due to contrasts among weekly means that do not reflect systematic training effects; the exact nature of these effects was not determined due to the huge number of post-hoc tests that would have to be conducted to find a significant but uninteresting contrast. The remaining two
significant measures, Appropriate Verbal (AV) behavior of targeted children and Inappropriate Physical (IP) behavior of non-targeted children reflect training effects, and are discussed in greater detail.

The observed significant decrease in IP of non-targeted children was found to have a significant linear component over trials, but no significant component for type of training. This decrease in IP behavior can be interpreted as a general training effect over time, but without the results of a no-treatment control group to compare with, other interpretations are possible. For example: inappropriate, disruptive physical classroom behaviors are highly visible to observers, and are probably embarrassing reminders to teachers that their classes are not well managed. It is plausible to postulate that as training progressed, teachers tried to eliminate such behaviors--using whatever methods they could--in response to the presence of observers. The fact that targeted children, who were by definition pupils with severe behavior problems, did not display a similar decrease in IP behavior is consistent with this interpretation. Inasmuch as teachers had identified targeted pupils for the observers, there may have been a greater expectancy that the observers would view the targeted pupils' IP behavior as "normal", and less of a cause for teacher embarrassment. Since targeted pupils were consensually defined as children who did not respond to usual classroom discipline, teachers may have been less motivated to try to reduce the IP behavior of targeted pupils for the benefit of the observers. The foregoing is speculative, but it does point to the need for a no-treatment control group to control for effects of observers on classroom routine.
The observed increase in AV behavior of targeted children shows a significant linear trend and a significant effect for type of training. In this instance, there is a somewhat stronger reason to interpret this finding as an actual training effect. While the linear trend could be attributed to a general "training program" effect or observer effects, the significant increases between phases (baseline, weeks 3-6, and weeks 7-10) strongly suggest an effect for type of training, even in the absence of a control group. Changes in AV behavior coincided with changes in type of training, and no significant changes in AV behavior were found within any given phase of training. It is reasonable to conclude, therefore, that with respect to AV behavior of targeted children, didactic training significantly increased this behavior over baseline rates and in-class practice of behavior modification skills further significantly increased this behavior of targeted children.

Overall, the results of the observational measures do not suggest a powerful training effect on teacher and pupil behavior; three-fourths of the measures showed no significant changes at all. On the other hand, there is no a priori reason to assume that the training program used in the present investigation would affect all pupil and teacher behaviors equally. Unfortunately, the literature provides no clues as to why the behaviors that evidenced change did so. There appears to be no logical relationship between the topics emphasized in the training program and the behaviors that changed significantly. It is possible that the coding system for observations was not sensitive to changes in teacher and pupil behavior, but this
is doubtful for two reasons. One, investigators using similar coding systems, e.g., Randolph and Wallin (1973), have found them sensitive to behavior change in the classroom. Secondly, the present author thinks that teachers cannot become successful modifiers of pupil behavior without making fairly gross changes in their patterns of attention—changes that would be reflected in the data, using the present coding system. Since the present author's thoughts may not reflect empirical reality, however, further research is, of course, indicated.

Correlations between pairs of teacher and pupil behaviors were calculated primarily to provide some insights into general patterns of interaction in the classrooms. Most of the significant correlations in Tables 9 and 10 are self-explanatory and reflect what one would expect to find in most classrooms. For example, in Table 9, teachers' Verbal Praise ($V^+$) correlates highly and positively with neutral physical touches and gestures ($P$), simply suggesting that teachers look at or are in close physical proximity to pupils when praising them. As another example, in Table 10, Inappropriate Physical ($IP$) behavior of non-targeted children correlates significantly and negatively with teachers' $P^+$, $V^+$, and $V$ behaviors, suggesting that teachers did not attend to IP behavior of non-targeted children. Interestingly, while the same correlations are negative for targeted children (Table 9), they are not significant, suggesting that teachers were less successful in ignoring IP of targeted children. This might also explain why targeted pupils displayed IP behavior; teachers maintained it with their attention.
Using the correlational data in isolation, it is not possible to determine functional relationships between any given pair of teacher and pupil behaviors; correlation does not imply functional relationship or causality. When interpreted in view of the method by which data were recorded, however, some correlations do suggest functional relationships. In the classrooms, observers recorded teacher behavior only if immediately antecedent pupil behavior could be identified, i.e., recording of teacher behavior was contingent on identifying and recording antecedent pupil behavior. For any pair of teacher-pupil behaviors, a functional relationship can be inferred if: (1) the pupil behavior changed substantially over training; (2) the teacher and pupil behavior correlate significantly; (3) the teacher behavior of the pair occurs most frequently in the observational record as a consequence of the pupil behavior of the pair. These criteria are met for two pairs of teacher and pupil behaviors, AV of targeted children and teachers' V behavior, and for AV of targeted children and teachers' V+ behaviors (see Table 9). In both instances the pupil behavior (AV) increased significantly over training, and correlates significantly and positively with both V and V+. In the observational record, approximately 90% of the targeted pupils' AV behaviors were followed by a V or V+. Approximately two thirds of teachers' V and V+ statements to targeted children were found to be consequences of pupils' AV behavior. Thus, it is reasonable to infer that the observed significant increase in Appropriate Verbal (AV) behavior of targeted children was due to teachers' contingent use of Verbal Praise and Verbal-neutral (V+, V) statements over the course of training.
The results of the questionnaire data indicate that subjects learned a significant amount of information about behavior modification over the course of training. It is not a particularly exciting finding, since one may assume that teachers can learn verbal information by attending lectures and reading printed material; it is gratifying, however, to find that the program teaches, at least at the verbal level, what it is intended to teach. This is, of course, a desirable first step in any training endeavor.

Within the noted restrictions of methodology, the findings of the present study suggest that didactic training was effective in fostering desirable changes in teacher and pupil behavior. In-class practice of behavior modification skills appears to have further increased the training effect with regard to behaviors noted above; in two instances, pupil behavior appears to be functionally related to changes in teacher attention. Negative results far outweigh positive results, however; the inevitable questions arise as to why results were not more dramatic, and how future programs should be constructed to achieve better results.

An obvious change in training strategy is to increase the length of training programs so that more time is available to develop teachers' behavior management skills. Of the training programs reviewed in Chapter 2, the longest is about twelve weeks in duration, with the average program lasting about eight weeks. This may simply be not enough time for teachers to learn a wide range of behavior modification skills. Lack of training time may not be the only undesirable aspect of short-term training programs. Shemberg et al. (1972) suggest that
the necessary changes in classroom routine which occur when teachers try out new management methods, may actually inhibit the desirable effects of incipient behavior modification programs. Learning to reinforce, ignore, shape and prompt behavior takes time—time that teachers normally devote to their usual, partially effective classroom management methods. Short-term training programs may end before teachers have completed the transition from old to new management methods, neither of which are probably working effectively during the transition period. Thus, pupil behavior may actually get worse until the teacher has established her new routine using behavioral methods of classroom management. Training programs should be long enough, therefore, to ensure that the classroom routine has stabilized and to ensure that teachers have become comfortable with their newly-learned behavior modification skills.

While duration of training may be a relevant factor in influencing teacher skill acquisition and pupil behavior change, a possibly more relevant factor is method of training. Didactic training programs rely heavily on programming stimuli (training materials) which are antecedents of hoped-for changes in teacher behavior; manipulation of rewarding consequences contingent on teachers' skill acquisition is, however, often neglected. Edgar (1972), having failed to find significant changes in the behavior of teachers who underwent six weeks of didactic training, concluded that trainers should practice what they preach, and devote more attention to reinforcing teachers for learning behavior management skills; training programs should be conducted along operant lines.
Reinforcing trainees' learning is not a novel idea (see Zimmerman and Zimmerman 1971), but the literature suggests that trainers do not use reinforcement as effectively as they might.

To use reinforcement effectively, reinforcement must be fairly immediate, and approximations to desired behavior management skills should be reinforced. Rather than delaying reinforcement until the end of a training program (as the Zimmermans did), teachers should have the opportunity to earn valued reinforcers as they learn each component of successful classroom management (pinpointing, recording, reinforcement, extinction, etc.). The choice of presumed reinforcers presents a problem, but not an insurmountable one. Money is an obvious choice. Teachers could be asked to pay a more-than-nominal fee for the training program, with the understanding that they can earn all or part of it back, contingent on their weekly performance. The possibility of being "paid" say, twenty dollars, cash on the barrelhead, each time a weekly assignment is completed, could act as a powerful incentive. (The present author uses such a "refund" method in his current clinical practice with parents of behavior-disordered children, and with clients on behavioral weight control programs. Subjectively, it appears to work quite well.) Graduate school credit is also a possible reinforcer. Hall's (1971) training program is offered for credit, as was the Zimmermans' program in 1971. To be maximally effective, teachers should be given frequent feedback as to exactly which behaviors learned will earn what amount of college credit. A token system could be useful in this respect. A third form of reinforcement
could be social approval and praise, used in conjunction with more tangible rewards (money and college credit). The reinforcing value of rewards could plausibly be strengthened if they are presented to individuals in front of the group, if weekly records of achievement are posted for the group to see, and if teachers are instructed to make favorable comments about each others' progress, and are reinforced for doing so by the trainer. The actual reinforcement procedures applicable to any given training program are, of course, best determined in view of limitations of time, money, and other resources available. As a general principle, however, it seems crucial that trainers adopt the position that teachers learn best under the same sorts of contingencies that their pupils do, and that trainers use operant methods to modify (train) teacher classroom management behavior.

The foregoing suggestions for breaking teacher training down into smaller, reinforceable units has implications for outcome research. Although the mastery of classroom behavior management is a function of learning numerous component skills and the methods used to train those skills, training programs tend to be evaluated as if "training in behavioral classroom management" were a unitary variable. In fact, however, any given training program consists of a large number of variables that can affect the outcome of training. Characteristics of trainers, actual training methods employed at various times during a program, and the amount of extra material trainers provide in response to teachers' questions could all be relevant factors affecting outcome. The author suggests that future
research in this area focus on the relative contributions of these and other variables that can effect the outcome of training. Such research should yield results that not only more clearly pinpoint the important factors contributing to the outcome of any given program, but which also permit more valid comparisons to be made between different programs.

The program reported in this dissertation also has some strengths that might profitably be incorporated into future programs. The emphasis on teachers actually trying out newly learned management methods appears to be a useful element of training which enhances lecture material and reading assignments. Conducting training sessions in the natural classroom environment may also foster generalization from training session to class session (see Walker and Buckley 1972). Finally, the observational coding system used in the present study appears to have utility in documenting possible functional relationships between teacher and pupil behavior, which is often not possible when using simple frequency counts and a repeated measures group design.
APPENDIX A

DEFINITIONS OF CODING CATEGORIES

**V+:** A verbalization is coded V+ if it indicates pleasure, approval, or satisfaction with a pupil's behavior. V+ may be coded on the basis of content or inflection or both.

**Examples:**
- Very good work, Bobby.
- Nice!
- I like the way you did that.
- WOW!
- You made it to the circle on time!

**V-:** A verbalization that indicates displeasure, disapproval, or dissatisfaction with a pupil's behavior. Any verbalization that would otherwise be a cue or prompt, but spoken in an angry tone of voice, and whose content indicates a behavior should cease.

**Examples:**
- Stop that right now! (a cue, but in an angry tone)
- That is a very bad thing you did, Bill.
- If I catch you doing that again, off to time out!
- You read that very poorly.

**V:** Other attending vocalizations that cannot be clearly coded V+ or V-.

**Examples:**
- Uh-huh.
- Good. (said in a flat tone of voice)
- Yes.
- You have finished, you may go.
- Do it again, Sally.

**Note:** Inflection is often a clue to a + or -. Facial expression is also a good clue.

**P+:** A physical gesture or movement or expression that indicates the sentiments of a V+.

**Examples:**
- A hug, a kiss.
- A pat on the back.
- A smile.
- A nod of the head, usually with a smile.
P-: Any physical gesture, expression, or contact that expresses the sentiments of a V-.
Examples:
A spank or a hit.
Dragging a child.
Restraint of a child.
A frown.
Shaking of the head side to side, indicating "No".
Glaring at a child.

P: A physical gesture, expression or contact that cannot be coded as a P+ or P-.
Examples:
Teacher rests hand on shoulder of child at desk.
Teacher makes eye contact with pupil without smiling or frowning.
Teacher bends over student, or walks up to student.

AV: Any verbal response of a pupil that precedes teacher attention and is appropriate to the ongoing interaction. All verbal responses are coded AV unless clearly IV (see below).
Examples:
Pupil reading aloud.
Singing in a group.
Answering questions, correctly or incorrectly.
Asking questions. (unless teacher has indicated this is not appropriate)
Appropriate laughter, as in response to a story.
Peer verbal teaching.
Pupil praising pupil.

IV: Any verbal behavior that is inappropriate to or disruptive of ongoing interactions or work; any verbal response that the teacher has previously told the pupil is inappropriate.
Examples:
Cursing, shouting, screaming.
Whining.
Whispering to another pupil during quiet period.
Tattling, telling lies.

Note: A wrong answer to a question is not coded IV unless it is an obvious lie. The key to IV behavior is that it is disruptive, or that it is in opposition to a rule or statement previously made by the teacher.

AP: A non-verbal behavior that the teacher has indicated before hand as being appropriate; a non-verbal behavior that is not clearly IP (see below).
Examples:
Complying to a request for a motor response.
Sitting at desk, standing in line, walking not running.
Sharing materials, cleaning up.
Listening quietly, handraising.
**IP:** A disruptive physical (non-verbal) behavior; a physical behavior previously defined as inappropriate by the teacher.  
**Examples:**  
Fighting, hitting, pushing.  
Throwing objects in class.  
Running in class.  
Nose picking, mouthing of hands or objects.  
Leaving seat, after being told to stay.

**G:** This code is used when the antecedent pupil behavior comes from more than one pupil. This code assumes no positive or negative value.  
**Examples:**  
Group responses to questions.  
Group singing.  
Group activities such as ring-around-Rosie.  
Any work or studying done by more than one pupil.

**Note:** The teacher will often say something that indicates antecedent pupil behavior was G. A pupil engaged in group behavior may be singled-out by the teacher for his behavior. In such instances, use the appropriate individual code. If the teacher praises or reprimands pupils engaged in the same activity and names each one, code as G.
The first weekly session began with a lecture covering the basics of social learning theory and behavior modification. Topics included: advantages of changing behavior in the classroom as opposed to referring the child elsewhere; the necessity of stating problems in terms of overt behavior; pinpointing behavioral objectives; the use of rewards as an alternative to changing behavior by the use of punishment. Trainers staged a skit portraying one child hitting another. Teachers were asked to describe what they saw, and were praised for specifying behaviors as opposed to using words like "bad" and "aggressive". The assignment for the first week was to make a list of ten appropriate behaviors and ten inappropriate behaviors pupils displayed in the classroom.

Lecture topics the second week were: using reinforcement; identifying reinforcers in the classroom; data collection using frequency counts and interval samples. Trainers presented a skit about a child who did not comply to parental commands. Teachers observed and recorded the frequency of non-compliance, using an operational definition supplied by the trainers. Assignments were: listing reinforcers available in the classroom; identifying personal-social reinforcers (hugs, smiles, praise) already in use in the classroom; observing and recording one "good" and one "bad" behavior for one week.
During the third week's training session, the concept of contingencies was introduced, and a distinction drawn between reinforcement and reward. Rules for making reinforcement work effectively (e.g., immediacy, consistency) were discussed, as was the use of time-out to decelerate behavior problems. Skits were presented showing the correct and incorrect use of reinforcement. The assignment for the third week was to formulate a tentative contingency-management plan for one pupil.

The application of timeout was the topic of the fourth session. The correct use of timeout was presented in a skit, and suggestions were made as to what constitutes a good timeout area. Assignments included continuing to record behaviors in class and finalizing the tentative contingency management plan. Trainers helped subjects with the latter assignment when teachers met in small groups after the lecture.

Shaping was discussed during the fifth lecture, and examples were presented as to how shaping could be used to train attention, non-hesitant oral reading, and cooperative play behaviors. Modeling appropriate behavior was briefly discussed, as was the use of intermittent schedules of reinforcement to maintain behavior. Trainees implemented their contingency-management plans during the fifth week. Other assignments included: developing a shaping program, whether or not it would be used; continuing to record data in class; revising contingency-management plans as needed.
The use of rules in the classroom was the major topic for the sixth week. Teachers were told that rules should cover important behaviors only, should be clear and enforceable, and should specify consequences of being followed. The role of peer reinforcement in maintaining inappropriate behavior was discussed, and suggestions were given regarding how to eliminate this problem, e.g., by praising class members for not attending to the deviant behavior of a peer. Teachers were assigned the task of developing a contingency-management plan for the entire classroom, incorporating suggestions about rules and peer reinforcement.

The seventh session consisted primarily of a discussion about the advantages and disadvantages of punishment. The trainers and authors of the manual had a definite bias against instructing teachers to use forms of punishment other than timeout. Thus, the discussion of punishment was mostly a "pitch" for the use of timeout and reinforcement of incompatible behavior. Side effects of punishment—avoidance, anxiety, and unintended reinforcement of inappropriate behavior were also discussed. No new assignments were made, but considerable time in the small groups was devoted to reviewing and revising existing interventions.

During the eighth and final week of training, the trainers presented a model for solving problems in classroom management. The model consisted of five steps: (1) problem identification and specification; (2) problem documentation, which consisted of a review of pinpointing behaviors and data collection; (3) brainstorming solutions; (4) picking and implementing a probable solution; (5) verification of
results of the chosen solution. The session ended with a discussion entitled "Arming Oneself Against Skeptics". Trainers presented and answered common questions and misconceptions about behavior modification, addressing such issues as reinforcement versus bribery, training versus "cold-blooded manipulation", and teachers' worries that children who work for extrinsic reinforcement will become materialistic and lose self-motivation. The author assumes that readers are familiar with such issues, and will not clarify them here. For a discussion of these and other issues in behavior modification, the reader is referred to About Behaviorism (Skinner 1974). The final assignment was for teachers to write descriptions of all interventions attempted, and to return them, with supporting data, the following week.
APPENDIX C

TEACHER QUESTIONNAIRE

Name________________________

Years Teaching Experience________

This questionnaire is designed to find out about your knowledge of topics that will be covered in the training program. Your responses to this questionnaire will help us in evaluating the program.

Please answer the multiple-choice items by darkening the "0" next to the phrase that best completes the item.

1. Behavior is most influenced by _____.
   0 one's heredity
   0 immediate consequences
   0 one's personality
   0 delayed consequences

2. If you want to decrease a behavior, _____the reinforcers maintaining it.
   0 gradually decrease
   0 slightly increase
   0 eliminate entirely

3. There are two kinds of behavior consequences.
   0 rewarding and punishing
   0 objective and subjective
   0 physical and verbal
4. In order to have the most impact on behavior, consequences
   0 must be rewarding
   0 must be unidirectional
   0 must be immediate

5. If you wish to increase a behavior that is just being learned,
   0 reward the behavior occasionally
   0 reward the behavior every other time
   0 reward the behavior every time

6. A rewarding consequence ______ the rate of the behavior it immediately follows.
   0 increases
   0 decreases
   0 sometimes has no effect on

7. To maintain a behavior that is already learned,
   0 reward it occasionally
   0 reward it every other time
   0 reward it every time

8. If you wish to quickly decrease a behavior
   0 ignore it consistently
   0 use logical consequences consistently
   0 use timeout consistently

9. Some of the side effects of punishment are
   0 anxiety and avoidance
   0 decrease in self-concept
   0 physical and psychological arousal
   0 all of the above

10. If you wish to quickly eliminate a self-injurious behavior
    0 ignore it consistently
    0 punish it consistently
    0 reward incompatible responses consistently
11. An event occurring immediately before a behavior is called
   0 a consequence
   0 a DRO
   0 a cue

12. Using a high-probability behavior to reinforce a low-probability behavior is called the
   0 Peter principle
   0 Premack principle
   0 Skinner principle

13. A reinforcer is defined by ________.
   0 its effect on behavior
   0 its intrinsic value
   0 its effect on the nervous system
   0 its extrinsic value

14. Which of the following is most true?
   0 nagging cannot be a reinforcer
   0 spanking cannot be a reinforcer
   0 yelling cannot be a reinforcer
   0 none of the above can be a reinforcer
   0 all of the above can be reinforcers

15. Pinpointing is the process of ________.
   0 accurate empathy
   0 accurate definition
   0 accurate reinforcement

16. A contingency is a (an) ________ relationship.
   0 upon...therefore
   0 either...or
   0 sometimes...always
   0 if...then
17. A negative reinforcer is \[_______\].
   0 roughly the same as punishment
   0 the termination of an unpleasant event
   0 roughly the same as a logical consequence

18. Tokens should be \[_______\].
   0 paired with social reinforcement
   0 used alone to increase their discriminative value
   0 paired with negative reinforcement
   0 all of the above

19. A powerful way to decrease a behavior is to \[_______\].
   0 reward incompatible behaviors
   0 reward compatible behaviors
   0 reward related behaviors

20. Cues are most often used to \[_______\].
   0 decrease behaviors
   0 start behaviors
   0 remember behaviors

21. A baseline is \[_______\].
   0 similar to a reinforcer
   0 a period of observation only
   0 the discriminative function of a verbal cue
   0 none of the above

22. Combining praise and punishment at the same time will probably
   0 be more effective than either one alone
   0 be less effective than either one alone
   0 be about as effective as using either one alone

23. Timeout is best characterized as \[_______\].
   0 a small room
   0 a dull, boring room
   0 a dark room
   0 a pleasant but quiet room
24. "Think small" should be the rule when using ________.
   0 social reinforcement
   0 shaping
   0 prompting
   0 physical reinforcement

25. Rules can be classified as ________.
   0 statements about contingencies
   0 statements about behavior
   0 statements about reinforcement

26. A behavior problem may be identified by ________.
   0 its high frequency
   0 its low frequency
   0 its setting
   0 all of the above
   0 high and low frequency but not setting

27. Behavior modification
   0 is a communist plot
   0 is based on a descriptive, empirical learning theory
   0 is based on an analytic, dynamic learning theory
   0 is based on the work of Pavlov, primarily
   0 is an extension of the theory of logical consequences
REFERENCES


