A MULTIPLE MEASUREMENT OUTCOME EVALUATION
OF A BEHAVIORALLY ORIENTED INSERVICE TRAINING PROGRAM

by

DOUGLAS SPENCER LEE
B.A., THE UNIVERSITY OF BRITISH COLUMBIA, 1974
M.Ed., THE UNIVERSITY OF BRITISH COLUMBIA, 1978

A THESIS SUBMITTED FOR PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES
(Department of Psychology)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
Feb. 1983

© DOUGLAS SPENCER LEE, 1983
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my department or by his or her representative. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

signed ____________________________

Department of ____________

The University of British Columbia
2075 Wesbrook Place
Vancouver, B.C.
Canada
V6T 1W5

Date ____________ FEB 21, 1983
This study investigated the immediate and long term impact of a portable, behaviorally-oriented, inservice training program. The training program consisted of a site visit (pre-test), 5 day workshop (mid-test), 5 follow-up visits, and a termination visit (post-test). The termination visit occurred 6-8 weeks following the last follow-up. Each training program spanned approximately 8-9 months. The training package consisted of lectures, films or video-tapes, demonstrations, role-play, behavioral rehearsal, and homework. Major areas covered during the training were teaching techniques (specifically Discrete Trial Format teaching), program writing, language training, curriculum development, reinforcement principles, maintenance and generalization, behavior management and token economy. The training team was composed of a coordinator, psychologist, teacher and community nurse.

Thirty-two direct service staff for severely handicapped clients served as trainees. These trainees were from training programs conducted in four communities in the Province of British Columbia from 1979-1981.

Data on three dependent measures plus consumer satisfaction measures were gathered using a pre-mid-post institutional cycle design. The three dependent measures were behavioral terminology, Discrete Trial Format teaching and program writing. Consumer satisfaction data were collected at the completion of the workshop and at the termination visit.

For the terminology and Discrete Trial Format teaching measures significant increases were found at mid-test and post-test in the trainees' scores and in the proportion of trainees exceeding pre-
established criteria for these measures. No pre-test program writing data were available. There was a significant decrease in the program writing scores from mid to post-test. Consumer satisfaction data was uniformly positive at both measurement periods.

The significance of these results are discussed in relation to previous research done in this area. Inferences are made about how factors such as trainee employment contracts, teacher schedules, and reinforcement schedules may have influenced the results. Problems in conducting this type of research under field conditions are identified and potential areas for further research are suggested.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>viii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>LITERATURE REVIEW</td>
<td>10</td>
</tr>
<tr>
<td>Lecture vs. Roleplay</td>
<td>15</td>
</tr>
<tr>
<td>Lecture vs. Lecture and Modeling</td>
<td>15</td>
</tr>
<tr>
<td>Modeling vs. Modeling and Lecture</td>
<td>18</td>
</tr>
<tr>
<td>Programmatic Packages</td>
<td>20</td>
</tr>
<tr>
<td>Summary of Review</td>
<td>28</td>
</tr>
<tr>
<td>HYPOTHESES</td>
<td>31</td>
</tr>
<tr>
<td>METHOD</td>
<td>33</td>
</tr>
<tr>
<td>Selection of Subjects</td>
<td>33</td>
</tr>
<tr>
<td>Trainees</td>
<td>33</td>
</tr>
<tr>
<td>Target clients of trainees</td>
<td>33</td>
</tr>
<tr>
<td>Settings</td>
<td>36</td>
</tr>
<tr>
<td>Description of Training Team</td>
<td>36</td>
</tr>
<tr>
<td>Training Program Description</td>
<td>41</td>
</tr>
<tr>
<td>Site visit</td>
<td>41</td>
</tr>
<tr>
<td>Workshop</td>
<td>41</td>
</tr>
<tr>
<td>Follow-up visits</td>
<td>42</td>
</tr>
<tr>
<td>Design</td>
<td>44</td>
</tr>
</tbody>
</table>
Measures

Knowledge of terminology
Program writing
DTF teaching skill
Consumer satisfaction: workshop feedback
Consumer satisfaction: follow-up feedback
Reliability

RESULTS

Knowledge of Behavioral Terminology
Discrete Trial Format Teaching
Program Writing

Tests on Proportion of Trainees Meeting Criteria

Proportion of trainees exceeding terminology test criterion
Proportion of trainees exceeding DTF criterion
Proportion of trainees exceeding program writing criterion

Consumer Satisfaction

DISCUSSION

REFERENCE NOTES

REFERENCES

APPENDICES

APPENDIX A
APPENDIX B
APPENDIX C
APPENDIX D
APPENDIX E
APPENDIX F
APPENDIX G
APPENDIX H
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Training Programs: Experimental Designs, Populations, and Treatment Outcomes</td>
<td>11</td>
</tr>
<tr>
<td>2. Training Group Composition by Occupational Category of Trainee</td>
<td>34</td>
</tr>
<tr>
<td>3. Age, Sex and Diagnosis of Target Clients</td>
<td>35</td>
</tr>
<tr>
<td>4. Schedule of Workshop Training Components</td>
<td>38</td>
</tr>
<tr>
<td>5. Total Time Allotment of Training</td>
<td>43</td>
</tr>
<tr>
<td>6. Contingency Tables for Computing Phi Coefficient for the Reliability of the Three Dependent Measures</td>
<td>50</td>
</tr>
<tr>
<td>7. Size of Included and Excluded Groups and Group Mean Scores for Workshop Participants on Dependent Measures</td>
<td>53</td>
</tr>
<tr>
<td>8. Mean Percentage Correct on Terminology Measure</td>
<td>55</td>
</tr>
<tr>
<td>9. ANOVA on Terminology Data</td>
<td>56</td>
</tr>
<tr>
<td>10. Mean Percentage Correct on DTF Measure</td>
<td>57</td>
</tr>
<tr>
<td>11. ANOVA on DTF Data</td>
<td>58</td>
</tr>
<tr>
<td>12. Mean Percentage Correct on Program Writing Measure</td>
<td>62</td>
</tr>
<tr>
<td>13. ANOVA on Program Writing Data</td>
<td>63</td>
</tr>
<tr>
<td>14. Contingency Tables of Trainees Meeting 80% Pass Criterion for Terminology Measure</td>
<td>65</td>
</tr>
<tr>
<td>15. Contingency Tables of Trainees Meeting 80% Pass Criterion for DTF Measure</td>
<td>66</td>
</tr>
<tr>
<td>16. Contingency Table of Trainees Meeting 70% Pass Criterion on Program Writing Measure</td>
<td>67</td>
</tr>
<tr>
<td>17. Mean Consumer Satisfaction Ratings for the Workshop and Follow-up Visits</td>
<td>69</td>
</tr>
<tr>
<td>18. ANOVA's on Workshop and Follow-up Consumer Satisfaction Data</td>
<td>70</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

FIGURE                  PAGE
1. The Institutional Cycle Design Used to Study the Impact of the Provincial Team's Training Program  45
2. Interaction of Group Mean Scores on DTF Measure  59
ACKNOWLEDGEMENTS

The hard work and generosity of many people made this project possible. I am deeply grateful to my advisor, Ken Craig, and thesis committee members, Bob McMahon and Susan Butt for their guidance and feedback which helped direct me to this final product. My sincerest thanks go to Shari Bender, Larry Bierbach, Dot Ewen, Jane Kelly, Karl Perrin, Barbara Thomas-Bruzzese, and Debbie Wilson, members of the Provincial Inservice Resource Teams without whose dedicated work this research would not have been possible. Finally, my love to Pat and the girls, Carol, Dianne and Monica for putting up with me when I was either staring blankly at walls or running into them.
INTRODUCTION

The history of theoretical concepts and systematic practices concerning care and training for the mentally retarded dates back to the early nineteenth century. This history, which has been well documented by Crisen (1975), Kanner (1967), and Rosen, Clark and Kivitz (1976) can be divided into four relatively distinct phases.

1) 1800-1850

The pioneering work was begun by Itard in France, Sequin in France and the United States, Guggenbuhl in Switzerland and Howe in the United States. Due to the humane and enlightened care given the mentally retarded during this period it became labelled the moral treatment era. Itard and Guggenbuhl both initially conceived of mental deficiencies as curable. They saw mental deficiencies as stemming largely from the person's lack of exposure to needed levels of sensory stimulation. Following this concept they devised sensory stimulation programs and sensory training programs within general education systems. The development of programs of this type led Itard and Guggenbuhl to develop special teaching techniques, in order to give their pupils the needed sensory training.

Although Itard and Guggenbuhl became disheartened by their inability to "cure" their pupils the demonstrable progress of their pupils greatly impressed the medical and scientific communities of their time. Sequin and Howe continued the pioneering work of developing educational strategies for the mentally retarded. Although Sequin endorsed the concept that mental retardation could be cured, Howe proposed a more moderate approach focused on the amelioration of the handicap.

2) 1850-1900

The early isolated programs of these four pioneers prompted a
positive attitude towards the teaching of the mentally retarded. The treatment facilities started by these pioneers were small and oriented towards each individual student. This initial concept of small and individually oriented resources was short lived, as many of the facilities quickly expanded and became overcrowded. The rapid growth of these small centers into large institutions during this time can be exemplified by their development in the United States. Prior to Howe's work in the mid 1800's there were no facilities in the United States for the mentally retarded. Howe started his residence in 1848. Over the next 40 years 15 facilities were opened. By 1898 twenty-four public institutions were maintained by 19 states and one by the City of New York. By 1917 all but four states were making some institutional provision. In total, approximately 37,200 patients were being cared for.

3) 1900-1950

The generally positive attitude about the nature of mental retardation, which was characteristic of the early 1800's, had shifted by the turn of the century. This shift culminated in a movement based on Galton's concept of eugenics, that was antagonistic to the well being of the mentally retarded. The most harmful interpretation of Galton's concept saw the unchecked fertility of the "unfit" (e.g., the mentally retarded) as a threat to the continuation of the positive inherited characteristics of the human race. This new movement was very active through to the 1930's and promoted a transformation of the original positive concepts of retardation and the type of care they should recieve. The new concept of retardation inappropriately linked the mentally retarded to a wide range of social maladies such as prostitution, robbery and vagrancy. It was also suggested that mentally retarded parents were the major and growing source of mentally retarded
offspring. Following this shift in thought, institutions increasingly became custodial care facilities more for the protection of society than the care and education of their residents. Although this movement lost its force in the 1940's, as new studies in genetics provided evidence contradictory to the eugenics concept, the attitudes and services which developed during the early nineteen hundreds persisted into the late 1950's.

4) 1950 - present

A number of forces started to surface during the 1950's which challenged the established beliefs regarding the care and treatment of the mentally retarded. The development of individualized intelligence testing, which started in the early 1900's, had yielded test data which led to viewing retardation as a gradient and not as a uniform syndrome. In conjunction with this change, specialized syndromes such as autism (Kanner, 1943) were isolated. DSM III (1980) now clearly distinguishes among four levels of retardation and among retardation, autism, and other pervasive or specific developmental disorders. From the 1930's onward accumulating evidence suggested that IQ's were not fixed but variable. Researchers also started to suggest that besides genetic factors, there were other factors such as poor nutrition, inadequate health care, and premature birth which might be significant as etiological factors in determining a person's level of intelligence.

Among with the changing scientific perspectives to mental retardation changes in attitudes towards the mentally retarded were occurring in society as well. Coinciding with the growing civil rights movements, a major force towards better treatment for the mentally handicapped (e.g., the mentally retarded, autistic and developmentally disabled) was the development of parents' associations. For example, the National Association for Retarded Citizens in the United States began as a small
group in 1950 but had approximately 50,000 members by 1959. The appointment of the President's Panel on Mental Retardation and their report in 1962 signified a major shift away from the concepts supporting institutional and custodial care for the mentally retarded. The report provided an outline of a national plan for services to the mentally retarded with a specific reference for the need to improve education programs and increase their availability. Thus, by the early 1960's a number of movements were well defined which, from various positions, were attacking the predominately negative attitudes and practices towards the handicapped. These attacks centered on the lack of training and remediation given to the retarded and the almost exclusive use of custodial care facilities as long term residences for this population.

All these developments led to a resurgence of a positive attitude towards, and more humane treatment for, the mentally handicapped. Two particular developments, originating in the early 1960's, have special relevance for the education and training of the mentally handicapped. The first development was the early demonstrations that operant techniques could be used to teach the mentally retarded new behaviors. The second development was the proposal of the normalization principle (Wolfensberger, 1972). The normalization principle states that the mentally handicapped should be offered patterns and conditions of everyday life which are as close as possible to the identified norms and patterns of the mainstream of society. Thus the normalization philosophy is the antithesis of a custodial care philosophy. The combination of the normalization philosophy with an operant technology, which provides teaching techniques for the mentally handicapped, has formed an educational movement for training the mentally handicapped to lead more independent and satisfying lives.
A major component of society's present commitment to the mentally handicapped deals with the rights of these individuals to an education. In the United States, court and legislative rulings, including the Pennsylvania Right to Education decision (1971), the Wyatt vs. Stickney decision in Alabama (1971), and Public Law 94-142 (1975), which legislated the right of all handicapped children to an adequate education, exemplify the present commitment (Martin, 1981). Although a different legal system prevails in Canada, a commitment to education for the mentally handicapped can be seen in the actions of several provincial education departments which have enacted legislation requiring the provision of educational opportunities for the mentally handicapped (Shepherd, 1982).

Fulfilling our responsibility to provide adequate education for the small proportion of the population which exhibits severely and profoundly retarded behavior (Grossman, 1977) or autistic behavior (Wing, Yeates, Brierley & Gould, 1976) is a major task. Kaufman (1981), dealing with the issue of whether all children were educable, listed eight conditions which make instruction of the severely mentally handicapped particularly difficult. His list follows:

1. All function at the extreme lower level of cognitive attainment and adaptive behavior.
2. Most have not acquired basic self-care skills.
3. Most are not known to many of the normal residents of their communities.
4. Most would never attend a public school unless the law specifically required their enrolment.
5. Most would have survived only a short time a few decades ago but have a life expectancy of decades due to modern medical intervention.
6. Most were, until recently, considered "untrainable" or "custodial" cases.
7. Most are non-ambulatory.
8. Most show extremely little promise of becoming creative, productive citizens even with the most heroic efforts of today's most skilled behavior therapists.

The autistic population has perceptual and cognitive impairments not found in the severely and profoundly retarded (Hermelin & O'Connor, 1970). The major behavioral indices of autism as listed in DSM III (1980) are: pervasive lack of responsiveness to other people, gross deficits in language development, and bizarre responses to various aspects of the environment. According to DSM III a differential diagnosis is possible between autism and retardation, although there is often substantial overlap between the two disabilities with 70% of autistic individuals having IQ's less than 70, and 40% with IQ's less than 50. DSM III allows for a joint diagnosis of both autism and mental retardation.

Attempts to deal with the above conditions, with a goal of providing adequate education for the severely and profoundly retarded and autistic populations, have led to the development of specialized teaching techniques (Koegel, Russo & Rincover, 1977; Liberty, Haring & Martin, 1981; Lindsley, 1964).

The development of operant teaching techniques for the severely mentally handicapped began in the late 1950's and early 1960's. Literature reviews on teaching techniques used with the severely mentally handicapped indicate that behaviorally based teaching techniques predominate within the field (Berkson & Landesman-Dwyer, 1977; Birnbrauer, 1976; Koegel, Egel & Dunlap, 1980; Lovaas & Koegel, 1972). In a review from 1977-1981 of four major journals specializing in mental retardation and autism (American Journal of Mental Deficiency; Journal of Autism and Developmental Disorders; Journal of Mental Deficiency Research; Mental Retardation) it was found that 86%
of the articles (n=42) specifically dealing with teaching techniques for the severely mentally handicapped populations dealt with behaviorally based teaching techniques. Treatments for autistic individuals have been formulated from psychodynamic, developmental, and physiologic arousal theories, but these have failed to gather empirical support (Lovaas, 1979; Lovaas & Koegel, 1972; Kazdin, 1975; Rimland, 1964).

The reviews referenced above indicated that, over the last 20 years, behaviorally based teaching techniques have been used for teaching a wide variety of skills to the mentally handicapped. This breadth of applicability continues to the present. Work in such areas as communication skills (Cuvo & Riva, 1980; Hung, 1980; Lovaas, 1977), self-help skills (Nelson, Gergenti & Hollander, 1980; Nutter & Reid, 1978; Schleien, Ash, Kierman & Wehman, 1981), vocational skills (Martin, Pallotta-Cunick, Johnstone & Goyos, 1980; Mithaug & Mar, 1980), socialization skills (Strain, Kerr & Ragland, 1979), independent recreational skills (Hill, Wehman & Horst, 1982), and gross motor skills (Gruber, Resser & Reid, 1979) indicate the variety of areas where behaviorally based teaching techniques have recently been used with the severely mentally handicapped.

Although behaviorally based techniques have been the treatment of choice to date it does not mean these techniques are without problems. Generalization and maintenance of teaching gains is a major concern (Rincover & Koegel, 1975; Koegel & Rincover, 1977). The large number of teaching trials, often in the thousands, needed to teach complex behaviors such as language (Goetz, Schuler & Sailor, 1979) may limit the practicality of teaching complex behaviors with the present technology. Finally, the technology has often not proven effective for teaching the client to emit a behavior with the proper latency, intensity or duration to allow the behavior to be easily
integrated by others into an ongoing social interaction (Liberty et al., 1981).

The amount of research evidence supporting the efficacy of behaviorally based teaching techniques is very encouraging. Yet, if the severely mentally handicapped are to receive maximum benefit from exposure to these techniques the direct service staff working with these populations must be skilled in their application of behavioral techniques. However, empirical evaluations of training programs designed to teach behaviorally based teaching techniques to direct service staff working with the severely mentally handicapped have been scarce (Donnellan, 1980; Gardner, 1972; Gladstone & Sherman, 1975; Panyon & Patterson, 1974).

The need for inservice training is particularly important at this time as pre-service training appears to be inadequate for preparing staff to work with these populations (Fredericks, Anderson & Baldwin, 1979; Haring, 1980). In the concluding remarks to their study attempting to define competency indicators of teachers of the moderately to profoundly retarded, Fredericks et. al., (1979) stated: "There is nothing in our experience, nor in the literature, to suggest that students who leave teacher-training institutions have the knowledge and experience to synthesize curricula, philosophy, and management techniques into a smoothly functioning, efficient, and successful educational system (p.93)."

The purpose of the present study was to measure the outcome of a highly portable inservice training program. This training program was conducted by a multidisciplinary team consisting of a psychologist, teacher, and community nurse. This team was trained by members of the Santa Barbara Autism Dissemination Project (Donnellan, LaVigna, Schuler & Woodward, Note 1) to conduct training programs using the Santa Barbara
model of in-service training for staff working directly with autistic and other severely mentally handicapped people. The Santa Barbara model used a 5 day workshop and a series of follow-up visits to ensure the maintenance and generalization of the techniques learned during the workshop to the trainees' natural working environment.
LITERATURE REVIEW

The intent of this review is to summarize research on the evaluation of training programs for direct service personnel who work with severely mentally handicapped populations. This review was conducted through the Psychological Abstracts from 1966-1981 and a direct search of journals specializing in the severely mentally handicapped and behavior modification. Several of the journals in these areas have only been published for brief periods of time. The journals specifically reviewed were: American Association for the Education of the Severely Profoundly Handicapped Review, 1979-1980; American Journal of Mental Deficiency, 1970-1982; Journal of Abnormal Child Psychology, 1974-1982; Journal of Applied Behavior Analysis, 1970-1982; Mental Retardation, 1970-1982. A total of nine articles was found which empirically investigated the effectiveness of programs for training staff to teach severely mentally handicapped clients. All nine programs focused on training staff to use behaviorally based teaching techniques.

The eleven investigations reported in these nine papers (Table 1) were categorized as attempts to research the effects of either isolated training components or the effects of treatment packages. As the majority of the studies focused on uniquely constructed and complex training methods as independent variables it was difficult to develop a systematic review. For example, Fredericks et al. (1979), Gardner (1972), Gladstone and Sherman (1975), Koegel, Glahn and Nieminen (1978), Panyon and Patterson (1974), Schinke and Wong (1977), and Weissman-Frisch, Crowell and Inman (1980) used lectures in their training packages. However, only Gardner (1972) and Panyon and Patterson (1974) focused specifically on the lecture component or indicated the duration of the lecture component. Furthermore, the other
### TABLE 1

Training Programs: Experimental Designs, Populations, Training and Results

<table>
<thead>
<tr>
<th>Authors</th>
<th>Trainees</th>
<th>N</th>
<th>Target Clients</th>
<th>N</th>
<th>Training Techniques</th>
<th>Research Design</th>
<th>Dependent Variables</th>
<th>Follow Up</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gardner attendant (1972)</td>
<td>2 groups of 10</td>
<td>20</td>
<td>retarded pop.</td>
<td>?</td>
<td>1. role-play 6 hrs</td>
<td>2 group counter-balanced</td>
<td>1. role-play play test</td>
<td>No</td>
<td>role-play train, incr. role-play proficiency: lectures incr. knowledge: order of pres. no diff.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. lectures 8 hrs</td>
<td></td>
<td>2. knowledge test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panyon &amp; attendant Patterson (1974)</td>
<td></td>
<td>3</td>
<td>retarded children</td>
<td>3</td>
<td>1. instruction multiple baseline</td>
<td>beh. obs. of reinf. tech. &amp; child % corr.</td>
<td>No</td>
<td>linear incr. of reinf. skill &amp; child % corr. across methods</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. VTR Feedbk.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. VTR model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gladstone high sch. &amp; Sherman stud. (1975)</td>
<td>14</td>
<td>7</td>
<td>profound retarded children</td>
<td>14</td>
<td>lectures, VTR model, behav. rehearsal &amp; correct feedback, praise</td>
<td>multiple baseline</td>
<td>beh. obs. of teaching skill</td>
<td>Yes</td>
<td>incr. skill in use of contingent reinf. &amp; instruction giving: pos. changes in child behs. results gen. to other child.</td>
</tr>
<tr>
<td>Authors</td>
<td>Trainees</td>
<td>N</td>
<td>Target Clients</td>
<td>N</td>
<td>Training Techniques</td>
<td>Research Design</td>
<td>Dependent Variables</td>
<td>Follow up</td>
<td>Results</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------------------</td>
<td>---</td>
<td>----------------</td>
<td>---</td>
<td>-------------------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>--------------------------</td>
<td>-----------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Schinke &amp; Wong (1977)</td>
<td>group-home staff of</td>
<td>30</td>
<td>retarded adults</td>
<td>?</td>
<td>lectures, case examples, role play, homework</td>
<td>2 group compar.</td>
<td>knowledge test, att. checklist, job satisfaction scale, beh. obs. of staff and residents</td>
<td>No</td>
<td>trained home (n=6) had sig. improve on att. check and pos. beh. change of staff and residents compared to control homes (n=6); both groups decr. in job satis, but trained home staff sig. less.</td>
</tr>
<tr>
<td>Fabry &amp; Reid (1978)</td>
<td>foster-grandparents</td>
<td>5</td>
<td>profound retarded adolescents</td>
<td>5</td>
<td>instruction, live model, prompts, praise</td>
<td>multiple baseline reversal</td>
<td>beh. obs. of teach interact. &amp; percent time not teaching</td>
<td>No</td>
<td>rate of teaching interact. incr. but only maint. for spec. beh.</td>
</tr>
<tr>
<td>Authors</td>
<td>Trainees</td>
<td>N</td>
<td>Target Clients N</td>
<td>Techniques</td>
<td>Research Design</td>
<td>Dependent Variables</td>
<td>Follow up</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------</td>
<td>---</td>
<td>------------------</td>
<td>------------------------------------------------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>-----------</td>
<td>---------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Koegel, Glahn &amp; Nieminen</td>
<td>mothers</td>
<td>4</td>
<td>autistic children</td>
<td>7</td>
<td>1. brief modeling demos.</td>
<td>multi-response</td>
<td>beh. obs. of teach. skill &amp; child's percent correct</td>
<td>No</td>
<td>increase in teach. skill and child's percent corr. consistent only after general training. mothers gen. across behs. and children.</td>
</tr>
<tr>
<td>Fredericks teachers</td>
<td>severely handicap. students</td>
<td>7</td>
<td>VTR modeling, beh. rehearsal with feedback</td>
<td>pre/post compari.</td>
<td>beh. obs. of Yes vari. teach. skills, rating of program</td>
<td>5 teachers met 7/7 crit- eria., 2 teachers met 6/7. teachers maint. skill at 1 yr. follow-up., pos. student gain found.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson &amp; Baldwin</td>
<td>therapists</td>
<td>3</td>
<td>autistic children</td>
<td>6</td>
<td>multi-response baseline</td>
<td>beh. obs.</td>
<td>No</td>
<td>increase in teach skill dependent on type of model viewed., child's perc. correct incr. only after the skill incr. in all areas.</td>
<td></td>
</tr>
<tr>
<td>(1978)</td>
<td></td>
<td></td>
<td></td>
<td>VTR model of antecedent teach behaviors</td>
<td>order of present.</td>
<td>percent correct.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1979)</td>
<td></td>
<td></td>
<td></td>
<td>VTR model of consequence teacher behavior</td>
<td>varied</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Trainees N</td>
<td>Target Clients N</td>
<td>Training Techniques</td>
<td>Research Design</td>
<td>Dependent Variables</td>
<td>Follow up</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>-----------</td>
<td>---------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weissman-Frisch, Crowell, &amp; Inman (1980)</td>
<td>direct service, administrator &amp; others 5-7 S's per group severe/profound adults</td>
<td>didactic, live &amp; VTR modeling, beh. rehab., with feedback, role play, homework. 4.5 day train program</td>
<td>pre/post compar.</td>
<td>beh. obs., rate scale., termin. test, goal attainment index, attitude scale re-test clients, consumer feedback, telephone interview.</td>
<td>Yes</td>
<td>consistent pos. increase across all measures from pre to post.: 10-12 week follow-up by tele. interv. 86-91% of S's indicated use of tech. from train program.: 60-64% of S's met their own goals.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
training programs with a lecture component had unique constellations of other components making up their full training program.

Although the literature review has been divided into four sections, the complexity of the training programs necessitated some overlap between categories. The four sections of this review examine different components of instructional methodology: lecture vs. roleplay; lecture vs. lecture and modeling; modeling vs. modeling and lectures; and programmatic packages.

**Lecture vs. Roleplay**

Gardner (1972) compared the use of direct versus indirect training methods for increasing institutional attendants' proficiency in the use of behavioral teaching techniques (giving instructions, prompting, shaping and data collection) and knowledge of behavioral principles. The training methods were six 1-hour roleplay sessions and eight 1-hour lectures. It was hypothesized that the direct methods (roleplay for increasing proficiency and lectures for increasing knowledge) would be more effective than the indirect methods (roleplay for increasing knowledge and lectures for increasing proficiency). The two groups of attendants (n=10) received the training methods in counterbalanced order. Regardless of the order of presentation roleplay was significantly better than lectures for increasing proficiency and lectures were significantly better than roleplay at increasing knowledge.

**Lecture vs. Lecture and Modeling**

Panyon and Patterson (1974) conducted two experiments which investigated the effects of modeling (live or videotaped) compared to videotaped playback of the trainees' own behavior, and instructions or lectures. As discussed by Bandura (1977), modeling is a very powerful method for teaching new behaviors, as it allows a person to develop symbolic representations of
the modeled behavior which can be used later as cognitive guides during the reproduction of the modeled behavior. In the first experiment a multiple baseline across three trainees (institutional attendants) was used. Two trainees received a training series of instructions in behavioral techniques (giving instructions and reinforcement), videotaped playback of the trainees own teaching behavior, and a final phase of live modeling. A third trainee received an instruction only phase followed by the live model training phase. The dependent measures were the trainees mean percentage correct for use of praise or verbal comment, primary reinforcement, and physical contact; and the percentage of correct or incorrect child responses. Due to the training phases not being counterbalanced for order of presentation it was not possible to evaluate the single training components. However, since trainee 3 did not receive the videotaped playback phase and still showed as large an increase in mean percentage correct as the other two trainees, the data suggested that the videotaped playback only phase was not required. The childrens' mean percent correct also linearly increased across phases.

In their second experiment Panyon and Patterson (1974) compared three training methods: 4 hours of lectures only, lectures plus film model, and lectures plus live model. A between groups design with three groups (n=5) of attendants as trainees was used. The dependent measures were the same as in their first study for the trainees' behavior. The results, analyzed using gain scores, indicated that both the lecture plus live model and lecture plus film model were equally successful and both were significantly better than the lecture only group at raising the trainees percent correct use of reinforcement. No data on the target childrens' behavior were given. The results from this study indicated that a lecture
or instruction plus model training format was effective for increasing the correct use of reinforcement.

Fabry and Reid (1978) used an initial instruction only phase in their training program of foster grandparents. However, in this phase, the instructions were about activities the foster grandparents could use to teach fine and gross motor skills to their mentally and physically handicapped clients. Following this initial stage, a training package involving live modeling, prompts, and praise was introduced to increase the foster grandparents proficiency in using behavioral teaching techniques (giving of instructions, prompts, and reinforcement). The mean percentage of intervals that grandparents engaged in teaching behaviors per session was the dependent measure. A multiple baseline design was used with the training for the grandparents introduced sequentially across three motor areas (head and neck movement, reaching, and manipulating objects). An increase in the percentage of intervals the grandparents engaged in teaching behaviors occurred only following the introduction of the training program across each of the three motor areas. Assessments of the residents' skill in fine and gross motor movements did not improve until the implementation of the training package for the foster grandparents had taken place.

These studies (Gardner, 1972; Fabry & Reid, 1978; Panyon & Patterson, 1974) indicated that, when used in isolation, lectures or instructions, either about specific behavioral techniques or about activities to use for teaching severely mentally handicapped clients, do not result in trainees increasing their teaching proficiency. Only when training techniques were introduced that allowed the trainees to roleplay or observe a model doing the teaching was proficiency increased. Positive changes in the target client's behavior
also did not increase until modeling of the teaching techniques was done for the trainees. However, lectures appeared to be more effective at increasing knowledge of behavioral principles than a roleplay training format.

**Modeling vs. Modeling and Lecture**

Koegel et al. (1978) conducted two experiments on the effects of modeling alone and modeling with lectures. The first experiment contrasted the effectiveness of a series of brief modeling demonstrations of a behavioral technique to a package of three 30-min. lectures describing the teaching technique in general, plus two 37-min. videotaped models, correctly and incorrectly using the components of the teaching technique. The behavioral teaching technique contained five components: $S^D$ (instruction or cue), prompts, shaping, consequences and discrete trials. Four parents served as trainees and the target clients were six autistic children. This experiment had two purposes; to assess the effectiveness of both formats for training the parents to teach a variety of target behaviors, and to assess if the parents could generalize their behavior across children.

Two dependent measures were taken. The first measure was based on the parents use of the five components of the teaching technique. The teaching sessions were broken into 30-sec. intervals and the parents' score for each component was computed for each 30-sec. interval, summed across all intervals and then averaged across all five components, to give an overall index of the parent's behavior per session. The second dependent measure involved computing the percentage of correct responses the children made to the parent's teaching.

Using a multiresponse baseline design, each of the parents received from one to four brief demonstrations before receiving the general training package. The results indicated that the parents maintained the proper use
of the technique only following the general training package. To obtain a high proficiency level with the teaching technique prior to receiving the general training package the parents needed brief demonstrations of the technique each time they attempted to teach a new target behavior. As all parents had at least one session of brief demonstration prior to the general training package it is not clear whether a general training only phase would have produced the same maintenance effects as a combined brief modeling then general training package.

During the brief demonstration phase the children's percentage of correct responding increased only following a demonstration for the parent of the correct procedure, focusing on the specific target behavior. Following the general training package the children's percentage of correct responding was maintained at a consistently high level, concurrent with the parents maintained high level of proficiency with the teaching technique.

The second experiment by Koegel et al. (1978) attempted to evaluate the effects of the two modeling videotapes used in their first study. One tape focused on the antecedent components of the teaching technique while the other focused on the consequent components. At the viewing of each tape the three subjects (teacher, foster mother, and undergraduate) were given a manual paraphrasing the script of the appropriate tape. The dependent measure was the same as in the previous experiment except the use of shaping was not measured and data were kept on the separate components of the teaching technique. Two subjects viewed the antecedent then consequence tapes, while the third subject viewed the tapes in the reverse order. Improvements in the correct application of the components were dependent upon viewing the appropriate tape, regardless of presentation order. This result indicated that training focusing specifically on the components to be taught, is necessary before proficiency in using those components is demonstrated.
Data on the percentage correct of unprompted child responses revealed that following the trainees' exposure to one modeling tape the children's percentage correct increased over baseline but the children's percentage correct was generally highest after the trainees were exposed to both films.

The results from Koegel et al. (1978) extend the information regarding the effects of modeling on teaching proficiency. Although brief focused modeling, lecture plus modeling packages, or video modeling of a generalized technique plus written scripts all increased teaching proficiency, only modeling of the technique under generalized conditions resulted in the trainees demonstrating maintained high levels of proficiency with the technique across different target behaviors and children. Brief focused modeling did not appear to provide the trainees with a learning condition which facilitated their development of a maintained proficiency with the teaching technique. The lecture and video modeling and video modeling with written script were equally effective at increasing trainee's proficiency, resulting in increased correct responding by children.

**Programmatic Packages**

The remaining studies all used multiple dependent measures and complex training packages as independent variables. Schinke and Wong (1977) used a training package of lectures, case examples, roleplays, weekly readings from the training book "Families" (Patterson, 1975), plus resident behavior change homework assignments to train group home staff in six homes for educable and trainable mentally retarded adults. Group home staff in six other homes provided a waiting list control group. The homes were randomly assigned to either the training or control conditions. The training took place in eight 1.5-hr. weekly sessions held at the group homes. The dependent measures were a knowledge test of behavioral concepts and techniques, an attitude
checklist of staffs' attitudes towards the residents, a job satisfaction rating scale, and naturalistic observations of 29 staff and resident behaviors. These 29 observation categories were collapsed for analysis into three general categories of positive, neutral, and negative behaviors.

As expected from such a training package there were significant pre-post gain score differences for the trained staff compared to the untrained staff on the knowledge test and on observations of the staffs' proficiency in adjusting their consequent behavior (positive, neutral, and negative) in relation to the residents' antecedent positive, neutral, and negative behaviors. Changes in overall rates of positive and negative behaviors in the trained homes were found. The trained staff showed significant increases in positive behaviors and significant decreases in negative behaviors compared to the untrained staff. Increased positive and decreased negative behavior rates of the residents were found concommitantly with the behavior changes of the trained staff. Significant positive increases in the trained staffs' attitude towards the residents were also found. An interesting finding was that job satisfaction ratings decreased pre to post for both groups, with the control group's decrease being significantly larger than the experimental group. Although the data on the knowledge test and naturalistic observations would appear consistent with the studies reviewed above, the use of multiple (37) univariate t-tests instead of multivariate statistics hampered a clear interpretation of the results.

Gladstone and Sherman (1975) trained seven volunteer high school students to work with profoundly retarded children. This study investigated the effects of a training package consisting of a 25-min. videotaped model,
behavior rehearsal, and corrective feedback every 5 minutes of rehearsal on the trainees' use of behavior modification techniques to teach profoundly retarded children. Following the initial training, which focused on one child, the trainees' ability to generalize their skills was assessed by having them teach a different behavior to a different child. A multiple baseline design across trainees was used.

The dependent measures used by Gladstone and Sherman (1975) were observations of the trainees' ability to give correct instructions, give reinforcers, not reward incorrect responses, use physical prompts, and ignore inappropriate and disruptive behavior. The childrens' percentage of correct responses was also used as a dependent measure.

Of the five dependent measures focusing on the trainees' behaviors only the use of instructions and contingent reinforcement showed consistent and substantial changes following training. The mean number of instructions given per session decreased by approximately 65% following training. The mean number of contingent reinforcers given per session increased by approximately 90% following training. These results were consistent across both target clients for each trainee, indicating that the trainees were able to generalize their behavior across children and behaviors.

The percentage of instructions followed by the children was variable. Four of the children in the initial training phase reached the 80% pass criterion while the other three children showed no improvement. However, during the generalization phase all children showed positive increases with six of the seven children reaching the 80% pass criterion.

An interesting point was that success in generalizing the teaching behavior to a second child was not dependent on the trainees' success with
teaching the first child. The three trainees who were unsuccessful in teaching their first child were able to teach their second child to criterion. However, the results on the ability of the trainees to generalize their behavior may be confounded, as only two target behaviors were used and they were not counterbalanced for order of presentation.

Koegel et al. (1977) investigated the effects of a programmatic package which focused on teaching behaviors similar to those studied by Gladstone and Sherman (1975). The experiment was designed to assess whether teachers could be taught to use behavioral teaching techniques correctly and whether correct use of the techniques had an effect on the target child's behavior. The training package consisted of a training manual, videotaped modeling, and behavior rehearsal with feedback every five minutes, with more elaborate feedback every 30 minutes. The training time was variable but all trainees completed the training within 25 hours. Eleven teachers were used as trainees and 12 autistic children acted as target clients.

The dependent measures focusing on teacher proficiency and child's percentage correct responding were the same as in the Koegel et al. (1978) study reviewed above. A modified multiresponse baseline design was used across trainees with the target children and behaviors changing for each session. The results indicated that the training package was effective in raising the teacher's level of proficiency to a consistently high level (over 80% correct). Measurement of the children's behavior revealed a positive effect of higher and more consistent correct responding following the teachers' completing the training package.

The results found by Koegel et al. (1977) indicated that the hypothesis that teachers use correct behavioral procedures and that the children can
not learn was incorrect. It appeared that, without training, teachers do not use behavioral teaching techniques properly and that under this condition the autistic children do not learn. However, when trained in the proper implementation of a behavioral teaching technique, teachers demonstrated a high level of proficiency with a resultant positive effect on the autistic child's performance.

The results from Gladstone and Sherman (1975) and Koegel et al. (1977, 1978) suggested two tentative conclusions regarding effective training. First, presenting the components of a teaching technique separately (Gladstone & Sherman, 1975) does not seem to result in as high a trainee proficiency as does presenting the components as a single technique (Koegel et al., 1977, 1978). Second, maintenance and generalization of trainee proficiency occurred only following a general training package.

Fredericks et al. (1979) conducted three experiments in an attempt to identify a set of skills, knowledge, and/or behaviors that could be trained to teachers and could be shown subsequently to affect student performance. The first experiment attempted to reveal behavioral indices which discriminated teachers of mentally handicapped children showing high learning gains from teachers of children showing low learning gains. The indices discriminating the high and low gain groups were: length of instruction day, percent of programs task-analyzed, training for aides, daily number of volunteers in class, seven items concerned with delivery of consequences and two items which combined cue and consequence delivery.

Following this first experiment, Fredericks et al. (1979) conducted an experiment in which a training program was developed to teach the identified behaviors to seven teachers from the low gain group. Using three classrooms at the Teaching Research Infant and Child Center in
Monmouth Oregon, the teachers, either individually or in groups of two, participated in a 5 day training program consisting of modeling, behavioral rehearsal, video and direct feedback, and seminars. An implementation manual outlining the areas covered during training was given to each trainee following the training program to facilitate the generalization of the techniques learned at the Center to their own classrooms. Training follow-up visits occurred approximately 4-5 months and 7-8 months following training. Eight performance objectives, which reflected the 13 behavioral indices found in the first experiment, were used to assess the teachers' proficiency at the completion of the 5 day training period.

The results on the eight performance objectives revealed that five of the seven teachers passed all criteria (pass criterion for the objectives ranged from 80%-90% correct) and the remaining two teachers passed seven of the eight criteria.

Data on four areas thought to be the most important from the 13 original behavioral indices were collected pre and post training and at a 12 month probe. The four areas assessed were delivery of consequences, percentage of time volunteers spend giving instruction, percentage of time instructions were delivered one to one, and percentage of programs task analyzed.

Data were collected in these four areas from all seven teachers at the three measurement points (except for one teacher whose administrator would not allow the use of volunteers in the classroom). Twenty-one of the 27 pre-post comparisons and 21 of the 27 pre-probe comparisons showed increases, with an overall analysis indicating that the pre-post and pre-probe gain scores for the four measures were all significant. Of the post-probe comparisons, 7 increased, 10 remained constant (all these post and probe scores were 100%) and 10 decreased. The mean decrease
was 21.8%, with the median decrease being 15%. Differences between the overall pre-post and pre-probe mean gain scores did not vary by more than three percent. It is instructive to note that the consistency in overall mean change scores plus the post-probe data would indicate that there was considerable variation in the trainees' post-to-probe scores with the impact of those trainees who increased their scores being balanced by those who decreased their scores. Therefore no reliable predictions could be made using individual teacher's post scores to predict their probe scores.

To test the effects of the training program on the students' behavior, the students (n=141) taught by trained teachers (n=14) were compared to an equal number of students of untrained teachers (n=14). Using the State of Oregon's mandated curriculum guide, average pre-post gain scores were computed for this comparison. The students in the classrooms of trained teachers gained significantly more than students of untrained teachers.

Weissman-Frisch et al. (1980) conducted a multiple perspective evaluation of an inservice training program, for vocational trainers of severely and profoundly retarded adults. The purpose of this study was to demonstrate the use of a multiple perspective evaluation approach to inservice training programs. This is a different purpose from the studies reviewed above, which focused on evaluating the effectiveness of a training method or package. This study examined an evaluation methodology. The key feature of a multiple perspective evaluation is that data are to be gathered from as wide a range of people (e.g., trainers, trainees, support staff or administrators) as possible, who are involved with the training program. In order to tap as many perspectives
as possible it was suggested that a much broader range of measurement devices is necessary than the traditional methods promoted in the literature.

Fifty trainees were given the 4½ day training program in groups of five to seven trainees. The three major content areas of the training program were task analysis, training methods, and production supervision techniques. The program package consisted of lectures, live and videotaped modeling, behavior rehearsal with feedback, and vicarious observations. The program was used at the Specialized Training Program at the University of Oregon. The Specialized Training Program contained a model work activity center which was used for several phases of the training program.

Dependent measures were developed for seven evaluation areas. Performance instruments were used to measure the trainees' skill in task analysis, use of training methods, and production supervision techniques. The Training Evaluation Kit was used for the trainees to provide consumer ratings of the training. The inservice instructors used the Training Decision Analysis Scale to weigh and rate attainment of inservice objectives. The Goal Attainment Index allowed trainees to weigh and rate specific objectives for each training component. A telephone interview was used as a follow-up to gather self-report data on the trainees' performance in the areas of task analysis, training methods, and production supervision. A survey regarding the trainees' attitude towards the mentally handicapped was given before and after training. Finally, a pre-post multiple choice test was used to assess knowledge of terminology and techniques.
Results from the multiple choice pre-test and self-report data stating a need for training indicated that from 50% to 92% of all trainees met the criteria for needing training in at least one of the three major content areas of the workshop. Post training results showed that 80% or more of the trainees passed both the multiple choice post-test and all performance criteria. Results regarding attitude change, consumer rating of the inservice program, and attainment of goals were all uniformly positive. The results from the attitude survey were of particular interest as the data revealed differential changes in attitude depending upon the level of mental handicap of the client being rated. Greater positive attitude change was noted for trainees rating their attitude towards the severely handicapped than the mildly handicapped.

From the telephone interview follow-up, 10-12 weeks post training, it was found that of the approximately 50% of trainees who wrote goals, between 60% and 64% of the goals relating to the three content areas had been met. Furthermore, the trainees indicated that they had trained 126 other staff in at least one of the three content areas of the training program.

The purpose of the Weissman-Frisch et al. (1980) study was to demonstrate a multiple perspective evaluation approach to an inservice training program. The range of measures used indicates the authors did fulfill their stated purpose of studying the breadth of impact an inservice training program may have.

**Summary of Review**

Several comments can be made with regards to all the studies reviewed. Institutional attendants, group home staff, parents, high school students,
foster grandparents, vocational trainers, and teachers are a very diverse group of trainees. Yet all the experiments reviewed were able to demonstrate that trainees showed increased proficiency in using behaviorally-oriented teaching techniques.

The trainees in these studies received their training from a variety of trainers. The types of trainers used were: behavior modification technicians (Gardner, 1972), attendants with more than one year of experience teaching the profoundly retarded (Panyon & Patterson, 1974), social work graduate students trained specifically for the study (Schinke & Wong, 1977), a university student (Koegel et al., 1977), a graduate student (Koegel et al., 1978), and a B.A. level psychologist (Fabry & Reid, 1978). It would appear then that the behaviorally-oriented teaching techniques used to teach severely mentally handicapped individuals can be taught by trainers with varying backgrounds to a wide array of trainees who work directly with these client populations.

The total duration of training given and the duration of training segments were factors which were highly variable across the studies reviewed. The total duration of training time varied from 1 hour (Koegel et al., 1978) to 5 days (Fredericks et al., 1979). The duration of training segments varied from 30-minute (Fabry & Reid, 1978; Koegel et al., 1978), to full working days (Fredericks et al., 1979; Weissman-Frisch et al., 1980). As would be expected there was a relationship between the duration of training and the number of content areas being taught. Fredericks et al. (1979) and Weissman-Frisch et al. (1980) had the longest duration of training and also covered the largest number of content areas. The other studies focused mostly on training specific
behavioral teaching skills and took less than 20 hours of training time, except the training by Koegel et al. (1977) which took 25-hours or less.

It has been demonstrated that training direct care staff to use behavioral teaching techniques has led to beneficial changes in the target client's behavior (Fabry & Reid, 1978; Fredericks et al., 1979; Gladstone & Sherman, 1975; Koegel et al., 1977, 1978; Panyon & Patterson, 1974; Schinke & Wong, 1977). The results from these studies, except for Fredericks et al. (1979), are limited to the time immediately following training. Fredericks et al. (1979) were able to demonstrate one year following training, that students of trained teachers made significantly greater increases in the number of behaviors learned than students of untrained teachers. Also, Fredericks et al. (1979) and Schinke and Wong (1977) were the only studies which indicated that the target clients' behavior of their trainees made significant changes within naturally occurring environments. Thus, while there is strong evidence that training staff in the proper utilization of behavioral methods produces immediate effects on target client behavior there is little evidence that these effects are enduring over time and under natural environmental conditions.
The purpose of the present investigation was to examine the effects of a training program on three areas of trainee behavior. These effects were tested under field conditions with the training offered on a cyclical basis to trainees in various communities throughout the Province of British Columbia. The three areas of trainee behavior that were measured were program writing, teaching technique, and knowledge of terminology. The program writing measure was an elaboration of the task analysis measurement used by Fredericks et al. (1979), while the teaching technique and measurement strategy were refinements of those introduced by Koegel et al. (1977, 1978). The technique introduced by Koegel et al. (1977, 1978) has become known as the Discrete Trial Format (DTF) of teaching (Donnellan, Gossage, LaVigna, Schuler & Traphage, 1977; Koegel et al., 1982). The terminology test was similar to the tests developed by Gardner (1972), Schinke and Wong (1977), and Weissman-Frisch et al. (1980) as a measure of the trainees' understanding of terminology related to behavior modification and the severely mentally handicapped.

At the completion of the workshop and at the final follow-up visit, consumer satisfaction measures were taken. Kazdin (1977) and Wolf (1978) have made strong arguments for the inclusion of consumer satisfaction measures as indices of the impact of clinical interventions. As suggested by Kazdin (1977), positive consumer satisfaction may be associated with maintenance of behavior change technicians given to clients.

Following the example of Fredericks et al. (1979) and Weissman-Frisch et al. (1980), who attempted to provide follow-up data on the effects of their training programs, the dependent variables of the present study were measured three times over an 8-9 month period. The hypotheses for this study were a) the trainees' scores on all three measures would be
higher at the completion of the workshop program (mid-test) and at the termination visit (post-test) than at the site visit (pre-test); and b) the proportion of trainees on each measure who scored above predetermined pass criterion set for each measure would be higher at the mid-test and post-test than at pre-test.
METHOD

Selection of Subjects

Trainees
All professional and paraprofessional personnel from four training programs conducted in British Columbia, who volunteered and completed the training program and were direct service providers for severely mentally handicapped clients, constituted the subject pool. Direct service providers were people who met the minimum requirement of working directly with clients on at least a once weekly basis. From a pool of 37 direct service providers, all those who completed all measurements on at least one dependent measure were included in this study. Thirty-two trainees (30 female and 2 male) met this criterion. Table 2 indicates the composition of the experimental training groups by occupational category. The mean years of education for all participants was 14.2 years with a range of 11-19 years.

Target clients of trainees
During the training program the trainees developed their skills while focusing on a single client before attempting to generalize their skills to all clients in their work setting. The target clients were chosen by the trainees. The target clients' learning capabilities (e.g. slow learning rate, lack of ability to generalize behaviors, poor prerequisite learning or work skill, or problems developing communication) were usually presenting great difficulty to the trainee. An alternative target client was chosen if, in the view of the team members, the proposed target client had such severely disruptive behaviors that immediate concern with managing their disruptive behaviors overshadowed the focus of teaching new behaviors. Table 3 lists the age, sex and diagnosis of
Table 2
Training Group Composition by Occupational Category of Trainee

<table>
<thead>
<tr>
<th>Training Group</th>
<th>Occupation</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Teacher</td>
<td>3</td>
</tr>
<tr>
<td>Vernon</td>
<td>Teacher Aide</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Nurse</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Psychologist</td>
<td>1</td>
</tr>
<tr>
<td>Group 2</td>
<td>Teacher</td>
<td>1</td>
</tr>
<tr>
<td>Nanaimo</td>
<td>Vocational Trainer</td>
<td>2</td>
</tr>
<tr>
<td>Group 3</td>
<td>Teacher</td>
<td>3</td>
</tr>
<tr>
<td>Williams Lake</td>
<td>Teacher Aide</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Childcare Worker</td>
<td>1</td>
</tr>
<tr>
<td>Group 4</td>
<td>Teacher</td>
<td>7</td>
</tr>
<tr>
<td>Kamloops</td>
<td>Teacher Aide</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 3
Age, Sex and Diagnosis\textsuperscript{a} of Target Clients

<table>
<thead>
<tr>
<th>Training Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Target Clients</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>11</td>
<td>6.7</td>
<td>14.7</td>
<td>11.8</td>
</tr>
<tr>
<td>Range</td>
<td>7-19</td>
<td>4-19</td>
<td>10-23</td>
<td>9-16</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Primary Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autism</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>- \textsuperscript{b}</td>
</tr>
<tr>
<td>Retardation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Severe/Profound</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Developmental-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delay</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Undefined Diagnosis</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

\textsuperscript{a}All specific diagnoses were from the clients' case records. Records for clients with undefined diagnosis either gave non-specific information about diagnosis or were not available to the training team. \textsuperscript{b}Individual diagnostic data for group four's target clients were unavailable. The intake criteria for the school that all clients attended were: IQ= 25-39, clients with IQ= 40-54 were accepted only if present functioning was seriously effected by social/emotional difficulties.
the target clients.

**Settings**

Training programs were conducted in four communities in British Columbia from June 1979 through to March 1981. The training locations, with the total number of participants and total number of participants who met the criteria of trainee listed in parentheses, were: Vernon (11,9), Nanaimo (7,3), Williams Lake (9,6), and Kamloops (16,14). In Vernon, Williams Lake, and Kamloops the workshop portion of the training program took place at the school where the largest number of target clients attended, while in Nanaimo a large meeting room adjacent to the school board offices was used. All follow-up visits were conducted at the work site of the trainee.

Chi-square analyses using the Yates correction for continuity (Ferguson, 1971) were computed to measure the distribution of occupational categories and educational levels of the participants across the four training programs. Occupational categories were teacher, teacher-aide, and non-school personnel. Education was divided into three levels, 11-13 years, 14-15 years, and 16 years or more of schooling. The results for occupational category were, $\chi^2 = 6.09, p = \text{n.s.}$ For education level the results were $\chi^2 = 3.56, p = \text{n.s.}$ These results indicate that although the training groups were comprised of trainees from a variety of occupations and educational levels, the group compositions did not significantly differ from each other on these two dimensions.

**Description of training team**

The purpose of the Provincial Inservice Resource Team was described in a government policy directive: "The purpose of the Provincial Inservice Resource Team is to train parents, teachers, and others to meet the needs of
children with severe behavioral disturbances, particularly those described as 'autistic' (including non-compliant, non-communicating children with severe developmental delays who show "autistic" behaviors). This service will support present residential or day treatment and classroom services by offering training to parents and staff responsible for autistic children in communities which do not, at present, have appropriate local services" (Noble, Note 2).

The Provincial Team consisted of a community nurse, teacher, psychologist and half-time coordinator (the coordinator was also responsible for the Woodlands Inservice Resource Team). The community nurse and psychologist were original members of the team who received their training from the Santa Barbara Autism Dissemination Project (Donnellan et al., Note 1), which conducted a training program for both the Provincial and Woodlands teams in 1978.

The contacts between the team and the users of the service could be broken into six components. These were: 1) request for service, 2) prereferral contact, 3) site visit, 4) 5 day workshop, 5) follow-up visits, 6) final follow-up visit or termination visit. The team's involvement with the users, from request for service to the termination visit, took approximately 8 to 9 months. The training program consisted of the site visit, workshop, follow-ups and termination visit, with the core of the training program being the workshop. Five follow-up visits were conducted, lasting from 2 to 4 days depending on the size of the training group. These follow-ups occurred initially at 2 to 3 week intervals, and faded to 4 or 5 week intervals between the last two visits. The termination visit occurred 6 to 8 weeks following the last follow-up visit.
Table 4
Schedule of Workshop Training Components

<table>
<thead>
<tr>
<th>Schedule of Inservice Components</th>
<th>Method of Training</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Day one</strong></td>
<td></td>
</tr>
<tr>
<td>A. Introductions to -</td>
<td></td>
</tr>
<tr>
<td>goals of the workshop</td>
<td>Didactic</td>
</tr>
<tr>
<td>autism</td>
<td>Didactic/Film</td>
</tr>
<tr>
<td>DTF</td>
<td>Didactic, VTR, Role-play, Demonstration</td>
</tr>
<tr>
<td>B. Practice with DTF teaching</td>
<td>Demonstration/Role-play</td>
</tr>
<tr>
<td>prerequisite skills and</td>
<td></td>
</tr>
<tr>
<td>teaching using criteria</td>
<td></td>
</tr>
<tr>
<td>C. Summary of DTF</td>
<td>Film</td>
</tr>
<tr>
<td>D. Readings</td>
<td>Homework</td>
</tr>
<tr>
<td><strong>Day two</strong></td>
<td></td>
</tr>
<tr>
<td>A. Review of readings and</td>
<td>Didactic/Discussion</td>
</tr>
<tr>
<td>introduction to baseline</td>
<td></td>
</tr>
<tr>
<td>B. Hands-on training with</td>
<td>Behavior rehearsal and</td>
</tr>
<tr>
<td>target clients</td>
<td>Feedback</td>
</tr>
<tr>
<td>C. Program Writing</td>
<td>Didactic, Behavior rehearsal and Feedback</td>
</tr>
<tr>
<td>D. Hands-on teaching using</td>
<td>Behavior rehearsal and</td>
</tr>
<tr>
<td>a written program</td>
<td>Feedback</td>
</tr>
<tr>
<td>E. Introduction to</td>
<td>Film</td>
</tr>
<tr>
<td>programmed curriculum</td>
<td>Didactic</td>
</tr>
<tr>
<td>F. Maintenance and generalization</td>
<td></td>
</tr>
<tr>
<td>G. Readings and writing program</td>
<td>Homework</td>
</tr>
</tbody>
</table>
Table 4 Cont.

Day three

A. Review and elaboration of program writing
B. Hands-on DTF teaching using a written program
C. Language and Communication
D. Hands-on teaching focusing on language training
E. Write language program

Day four

A. Principles of Behavior Management
B. Hands-on teaching using language programs for homework
C. Techniques of Behavior Management
D. Task analysis
E. Overview of last two days
F. Write a task analysis

Day five

A. Curriculum
B. Small group planning for the first follow-up visit
C. Review of homework
D. Token economy
| E. Review of trainees teaching using DTF | Model VTR segments of trainees teaching taken during workshop |
| F. Terminology mid-test | Knowledge of terminology test |
| G. Consumer satisfaction ratings | Workshop feedback form |
| H. Conclusion | Discussion |
Training Program Description

The present section contains an overview of the structure and components of the training program. A manual describing the conducting of a training program is contained in Appendix A.

Site Visit

The site visit was a preparatory visit which occurred 4 to 6 weeks prior to the workshop and lasted for 2 to 3 days. This visit had three major components: collection of pre-test data on terminology, program writing and DTF teaching measures; distribution of the training manuals (Donnellan et al., 1977); program text (LaVigna & Trophagen, 1978), readings, and finalization of necessary administrative details.

All three pre-test measures were introduced to the trainees as assessment methods the team used to tailor the content of the workshop to the strengths and weaknesses of the training group, as well as being used for a research project. Each trainee was individually videotaped while teaching the target child he or she would focus on during the workshop. The instructions for the teaching were that the teaching should be a 1-1 structured session of teaching the client a behavior he or she had not yet mastered. The trainees were asked to give the team a copy of the latest program they had written. The terminology test was given to each trainee with the instruction that the test was to be completed within 15 minutes.

Workshop

Table 4 presents the training components for each day and the methods of training used for each component. An example of the detailed workshop training schedule used for the training programs is contained in Appendix B. The order of introducing the major content areas was DTF teaching, program writing, maintenance and generalization, language/communication, behavior
principles and management, curriculum and token economy.

All didactic presentations, other than the introduction of goals for the workshop, were accompanied by articles or handouts. Trainees were instructed to bring a large binder to the workshop so all the articles and handouts could be collated to build a personal reference compendium. During hands-on training the trainees were videotaped on a rotating basis such that all trainees were videotaped at least twice during Days 2, 3, and 4. These tapes were reviewed and scored each night and verbal feedback was given each day regarding the progress the trainees were making in using DTF teaching. The programs and task analyses written for homework were collected at the beginning of each day. The programs were scored and written feedback was given on the program. Oral feedback to the group as a whole was given at the end of the morning session.

Each training day was 7 hours in duration. Excluding time for lunch and breaks, 5½ hours were available for training each day, or 27.5 hours for the whole week. Table 5 shows the total time allocated for the various training methods, for a total weekly training duration of 22 hours. The 5½ hours not totaled were used for concluding each day, giving out homework, planning for the first follow-up, completing the terminology test, and completing consumer satisfaction ratings.

Follow-up Visits

The major goal of the follow-up visits was to help the trainees integrate the skills learned during the workshop into their normal work environment. Once this was accomplished a second goal was to help the trainee generalize their newly developed skills to other clients. Five areas of interest formed the basis for most follow-up visits.
Table 5

<table>
<thead>
<tr>
<th>Training Methods</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Didactic/Discussion</td>
<td>10.5</td>
</tr>
<tr>
<td>Films/VTR</td>
<td>3.0</td>
</tr>
<tr>
<td>Role-play</td>
<td>1.5</td>
</tr>
<tr>
<td>Demonstrations</td>
<td>1.5</td>
</tr>
<tr>
<td>Behavior rehearsal</td>
<td>4.5</td>
</tr>
<tr>
<td>with feedback</td>
<td></td>
</tr>
<tr>
<td>General feedback</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22.0</strong></td>
</tr>
</tbody>
</table>
These areas were: observing the trainees teach, discussing any new written programs, reviewing data collection, discussing curriculum, and reviewing and monitoring behavior management programs.

All follow-up visits and the termination visit were conducted at the trainees' work setting. To provide links between follow-up visits goals were set for each trainee to work on between visits. Previous goals were reviewed each visit and new goals set for the next meeting. The periods between follow-ups lengthened from an initial 2 week period between the workshop and first follow-up to 5 or 6 weeks between the last two follow-ups. The termination visit, when post-test measures and follow-up consumer satisfaction ratings were taken, was 6 to 8 weeks after the final follow-up.

Design

Cook and Campbell (1979) used the term institutional cycle design to describe the design used for this study. This research design has been recommended by several authors (Campbell & McCormack, 1957; Campbell & Stanley, 1966; Cook & Campbell, 1979; McKillips, 1979; Powers & Alderman, 1979) for use when there is a cycle to the delivery of treatment or training. Figure 1 outlines the institutional cycle design as applied to the present study.

Campbell and Stanley (1966) have suggested that when the focus of field research is on the recurrent application of a complex independent variable, such as a training program, the institutional cycle design (termed the recurrent institutional cycle design by Campbell and Stanley) probably offers the best quasi-experimental design developed thus far. Arguments were also presented by Campbell and McCormack (1957) and Campbell and Stanley (1966) that if the program being offered is one
Figure 1

The Institutional Cycle Design used to Study the Impact of the Provincial Team's Training Program

1979 1980 1981

J JASONDJFMAMJJASONDJFMAMJ
u u a a u u
n n n n n n
e e e e e

Group 1
p m p
r i o
e d s t

Group 2
p m p
r i o
e d s t

Group 3
p m p
r i o
e d s t

Group 4
p m p
r i o
e d s t

As the mid-test occurred during the workshop the term mid indicates the mid measurement and the occurrence of the workshop.
which all potential subjects wish to have access to, then the
institutional cycle design may be better than a true experimental design.
The rationale for this possible superiority was that it may not be
feasible to find a proper control group for a true experimental design
which would allow for generalizations to be made to the natural
recruitment and training conditions.

The institutional cycle design is essentially a combination of the
longitudinal and cross-sectional approaches employed in developmental
research. As indicated by Campbell and Stanley (1966) and Cook and
Campbell (1979) the longitudinal and cross-sectional approaches
complement each other in controlling for threats to the internal and
external validity of the experiment. For example, a longitudinal design
does not control for history but a cross-sectional approach does. However,
a longitudinal design does control for attrition while a cross-sectional
approach does not. The most problematic threats to internal validity
with this design are testing, maturation and selection-maturation
interaction.

The institutional cycle design allows for several types of
comparisons. Each group can be looked at separately, as a pre-mid-post
case study of the training impact. Thus the data from each group can
be analyzed to see if they conform to the stated hypotheses. The groups
could be compared within a measurement period to assess the similarity
of the groups at a common phase of the design. This comparison could,
for instance, indicate differential effects of the training among the
groups at the mid-test measurement period. Finally, groups which are
measured at temporally similar times but are in different training phases
can be compared, to test for the effects of history. The cross-sectional
comparison of the mid-test for group 1 with the pre-test of group 2 where
the mid-test scores of group 1 are different than the pre-test scores
of group 1 provides differences which could not be explained by the effects of history.

**Measures**

Three performance measures plus two consumer satisfaction measures were collected. The three performance measures, knowledge of terminology, program writing and DTF teaching skill were obtained at the site visit, at the completion of the workshop, and at the termination visit. The consumer satisfaction measures were collected at the completion of the workshop and at the final follow-up.

**Knowledge of terminology**

The knowledge of terminology test was a 15 item combined multiple-choice and fill-in-the-blank test. It sampled items concerning behavior modification and issues related to training the severely mentally handicapped. It was used to assess trainees' knowledge of terminology used during the training program. An answer key ensured consistent scoring of the test. Appendix C contains a copy of the terminology test and scoring key.

**Program writing**

The aim of writing a program was to have a written document which indicated all the necessary information so that another person could read the program and duplicate the teaching environment and activities for the target client. The task analysis measurement used by Fredericks et al. (1979) and Weissman-Frisch et al. (1980) measured those parts of a program related to identifying all the subcomponents of a target behavior. A program contained not only a task analysis or series of teaching steps but also indicated the cue or instruction, conditions for using prompts, pass and fail criteria, specification of teaching
materials, schedule of reinforcement, how the target behavior was to be maintained or generalized, and auxiliary details necessary to ensure potential duplication of the training procedure.

A checklist of 40 items was used to score the written programs. A scoring guide which defined all items and gave examples of various items was used for all scoring. As several programs were collected from each trainee during the workshop, the highest score obtained was used as the score for the workshop. Appendix D contains a copy of the program writing checklist and the scoring guide.

**DTF teaching skill**

The DTF teaching measure was developed from the measure used by Koegel et al. (1977, 1978). The present version of the measurement device has six components: preparation, instructional stimulus, prompt, response, consequence, and intertrial interval. As each trainee was videotaped at least twice during the workshop the highest score obtained was used as the score for the workshop. Appendix E contains a copy of the DTF scoring sheet and scoring guide.

**Consumer Satisfaction: workshop feedback**

During the afternoon of the last day of the workshop trainees were asked to complete a feedback sheet containing items related to the organization and preparation of the workshop, content of the workshop and training methods used. Each item was rated on a 5 point scale. As each community had unique requirements some changes in content areas were made for each training program. An example of a workshop feedback sheet is contained in Appendix F.

**Consumer Satisfaction: follow-up feedback**

At the termination visit, trainees were given a two part feedback form. The first part asked questions regarding the trainees' use of the techniques learned during the training program. The second part asked
questions about how effective the team was in helping the trainees generalize and integrate the training materials into their work setting and whether their use of the techniques had helped any clients. All items which were rated used a 5 point scale. Appendix G contains both parts of the follow-up feedback form.

Reliability

Reliability checks were conducted to assess the consistency of interrater scores for the program writing and DTF teaching measures and to assess compliance in using the terminology answer key. The reliability coder for the program writing and DTF teaching measure was a member of the Woodlands Team. As the members from the Woodlands Team had never had contact with the Provincial Team's trainees this ensured blind ratings of these two measures. The terminology test reliability was determined by members of the Provincial team after all identifying information was concealed.

To ensure that reliability checks were conducted for each test period, the scores were grouped into the three measurement periods, and scores were then randomly selected from within time periods for the reliability check. The percentage of reliability checks averaged across the three periods were 39% for the terminology measure, 31% for the DTF teaching measure and 24.5% for the program writing measure. The lowest percentage of scores checked within any single time period was 21%. (Percentage of scores checked and percentage agreement per test period can be found in Appendix H).

The phi coefficient was used to calculate the reliability of the measures. Table 6 contains 2 X 2 tables indicating the frequencies of agreements and disagreements between the original coder and the reliability coder. The reliability is based on the agreements of the coders on an item level, rather than for total scores. In general this
Table 6
Contingency Tables for Computing Phi Coefficient
for the Reliability of the Three Dependent Measures

<table>
<thead>
<tr>
<th></th>
<th>Original Coder</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td>DTF</td>
<td></td>
<td></td>
<td>1180</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>33</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>1213</td>
<td>184</td>
<td>1397</td>
<td></td>
</tr>
<tr>
<td>Program Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>284</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>29</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>313</td>
<td>71</td>
<td>384</td>
<td></td>
</tr>
<tr>
<td>Terminology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>510</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Hit</td>
<td>Miss</td>
<td>13</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td>523</td>
<td>204</td>
<td>727</td>
<td></td>
</tr>
</tbody>
</table>
level of analysis provides a lower-bound estimate of reliability compared to reliability measured by total scores (Hartmann, 1977).

In addition to being a stronger statistical measure than percentage agreement the phi coefficient also allows for a test of the significance of the coefficient using chi-square (Glass & Stanley, 1970). The phi-coefficients were, for DTF teaching, \( \phi = 0.56 \), with \( \chi^2 = 438.38 \), \( p < .001 \); for program writing, \( \phi = .63 \), with \( \chi^2 = 152.39 \), \( p < .001 \); and for terminology test, \( \phi = .93 \), with \( \chi^2 = 623.12 \), \( p < .001 \). Thus all three measures were substantially and significantly reliable.
RESULTS

Initial analyses for homogeneity of variances were conducted using the F-max test (Winer, 1971). Significant heterogeneity of variances were found between test periods, on all three measures. An arc-sine transformation (Winer, 1971) reduced heterogeneity to non-significant levels. However, comparison of the results from analyses using the original and transformed data revealed only minor non-systematic differences between the results. Analyses from both types of data were in perfect agreement regarding significant versus non-significant results. The reason for the similarity of the results was that the sample size across test periods was equal, which allows for sizeable heterogeneity of variance without seriously biasing the results. As no major differences existed between the original and transformed data analysis it was decided to report all results using the original data.

To test for differences between those participants included in the analysis as trainees and those participants excluded from the analysis, the two groups were separated and the group mean scores at the mid-test were analyzed using t-tests. Table 7 indicates the number of participants included and excluded, as well as the group mean score for each measure. The comparison between the included and excluded group on the terminology measure indicated a non-significant difference, \( t(35) = .2837, p = \text{n.s.} \). For the DTF measure, the result was also non-significant, \( t(30) = .4069, p = \text{n.s.} \). For the program writing measure, the result was \( t(35) = 1.67, p = \text{n.s.} \), also non-significant. These results indicate that the scores of the included versus excluded participants were not significantly different at the completion of the workshop.
Table 7

Size of Included and Excluded Groups and Group Mean Scores for Workshop Participants on Dependent Measures

<table>
<thead>
<tr>
<th>Terminology</th>
<th>n</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included</td>
<td>29</td>
<td>83.0</td>
</tr>
<tr>
<td>Excluded</td>
<td>8</td>
<td>81.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DTF</th>
<th>n</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included</td>
<td>19</td>
<td>93.5</td>
</tr>
<tr>
<td>Excluded</td>
<td>13</td>
<td>92.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Program Writing</th>
<th>n</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included</td>
<td>15</td>
<td>89.87</td>
</tr>
<tr>
<td>Excluded</td>
<td>22</td>
<td>83.23</td>
</tr>
</tbody>
</table>
Knowledge of Behavioral Terminology

A 3 (test period) X 4 (training group) mixed effects ANOVA with repeated measures was used to test for differences in the terminology scores. Table 8 indicates the mean percentage correct for the terminology test and Table 9 summarizes the analysis of variance. A significant test period effect was found, $F(2,50) = 50.94$, $p < .001$. The groups did not differ, nor was there an interaction between groups and test period.

Post comparisons of the significant main effect were done using the Neuman-Keuls test for multiple comparisons on repeated measures (Winer, 1971). The harmonic mean of the cell sizes was used for these comparisons as the cell sizes were unequal. The significant comparisons were between pre-test and mid-test, $Q(2,50) = 24.66$, $p < .01$ and between pre-test and post-test, $Q(3,50) = 27.14$, $p < .01$. These results support the hypothesis that the trainees would score significantly higher at mid and post-test compared to their scores at pre-test on the terminology test.

Discrete Trial Format Teaching

A 3 (test period) X 4 (group) mixed effects ANOVA with repeated measures was used to test for differences in DTF teaching scores. Table 10 shows the mean percentage correct for DTF teaching measure and Table 11 presents the analysis of variance. Significant main effects were, Training Group, $F(3,15) = 5.54$, $p < .0092$; Test Period, $F(2,30) = 35.78$, $p < .0001$, and Training Group X Test Period interaction, $F(6,30) = 4.47$, $p < .0024$. Figure 2 demonstrates the interaction effect.

The interaction effect was broken down into simple main effects. The analysis of the simple main effect for training groups within each
Table 8

Mean Percentage Correct on Terminology Measure

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Pre</th>
<th>Mid</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>52.75</td>
<td>82.00</td>
<td>79.86</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>51.67</td>
<td>73.67</td>
<td>73.67</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>55.40</td>
<td>76.40</td>
<td>87.60</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>58.92</td>
<td>88.31</td>
<td>79.77</td>
</tr>
</tbody>
</table>
Table 9

Results of Analysis of Variance on Terminology Data

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>ms</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Group</td>
<td>3</td>
<td>242.417</td>
<td>0.99</td>
</tr>
<tr>
<td>Error</td>
<td>25</td>
<td>244.441</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Period</td>
<td>2</td>
<td>4707.842</td>
<td>50.94*</td>
</tr>
<tr>
<td>Period x Group</td>
<td>6</td>
<td>130.244</td>
<td>1.41</td>
</tr>
<tr>
<td>Error</td>
<td>50</td>
<td>92.419</td>
<td></td>
</tr>
</tbody>
</table>

*p < .0001
## Table 10

**Mean Percentage Correct on DTF Measure**

<table>
<thead>
<tr>
<th>Group</th>
<th>Test Period</th>
<th>n</th>
<th>Pre</th>
<th>Mid</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>7</td>
<td>36.00</td>
<td>92.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>43.50</td>
<td>94.00</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3</td>
<td>3</td>
<td>83.67</td>
<td>94.00</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>7</td>
<td>7</td>
<td>72.26</td>
<td>94.57</td>
</tr>
</tbody>
</table>
Table 11

Results of Analysis of Variance on DTF Data

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Group</td>
<td>3</td>
<td>925.605</td>
<td>5.54*</td>
</tr>
<tr>
<td>Error</td>
<td>15</td>
<td>167.139</td>
<td></td>
</tr>
<tr>
<td><strong>Within</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Period</td>
<td>2</td>
<td>6226.204</td>
<td>35.78***</td>
</tr>
<tr>
<td>Period x Group</td>
<td>6</td>
<td>777.213</td>
<td>4.47**</td>
</tr>
<tr>
<td>Error</td>
<td>30</td>
<td>174.010</td>
<td></td>
</tr>
</tbody>
</table>

* p < .01
** p < .005
*** p < .0001
Figure 2

Interaction of Group Mean Scores on DTF Measure

Mean Percent Correct

100
90
80
70
60
50
40
30
20
10

Group 3
Group 4
Group 2
Group 1

Pre Mid Post
Test Period
test period indicated a significant group effect within the pre-test period, $F(3, 15) = 4.67, p < .025$, but not during subsequent test periods. The analysis of the simple main effects for individual groups' test period revealed several significant effects. Test period was significant for group 1, $F(2, 48) = 43.29, p < .001$; group 2, $F(2, 8) = 10.98, p < .01$, and for group 4, $F(2, 48) = 13.26, p < .001$, but not for group 3.

Multiple comparisons were done on these significant simple main effects. To test the training group differences at the pre-test time, the harmonic mean for the four groups was used in the Neuman-Keuls test, as the groups had unequal cell sizes. The significant comparisons were between group 1 and group 3, $Q(4, 15) = 7.00, p < .005$; group 1 and group 4, $Q(3, 15) = 5.33, p < .005$; group 2 and group 3, $Q(3, 15) = 5.90, p < .005$, and between group 2 and group 4, $Q(2, 15) = 4.23, p < .01$. In other words, group 1 and group 2 were not significantly different from each other, group 3 and group 4 were also not significantly different from each other, but groups 1 and 2 both had pre-test scores significantly lower than groups 3 and 4.

The simple main effects for test period were further analyzed by multiple comparisons. The significant comparisons for group 1 were between pre-test and mid-test, $Q(2, 30) = 22.49, p < .005$, and pre-test and post-test, $Q(3, 30) = 23.23, p < .005$. For group 2, the significant differences were between, pre-test and mid-test, $Q(2, 30) = 10.84, p < .005$, and mid-test and post-test, $Q(3, 30) = 12.01, p < .005$. Significant differences for group 4 were between pre-test and post-test, $Q(3, 30) = 4.78, p < .01$, and pre-test and post-test, $Q(2, 30) = 4.14, p < .01$.

This pattern of results supports the hypothesis that the mid and post-test scores would be significantly higher than the pre-test scores, except
for group 3. As can be seen from Table 8, the pre-test mean for group 3 was already above the 80% criteria set as indicative of competency on this measure. This high pre-test score limited the possible amount of change towards supporting the hypothesis, even though there was an increase in mean scores at both mid and post-test periods.

Program writing

A 2 (test period) X 3 (group) mixed effects ANOVA with repeated measures was used to test for differences in program writing scores. The analysis was reduced in size for two reasons. First, as will be elaborated in the discussion section, only three subjects completed the pre-test measure which left insufficient pre-test data to make inclusion of data for this time period meaningful. Second, only one trainee in group 2 completed both the mid and post-test measures which precluded inclusion of group 2 in this analysis. Table 12 indicates the mean percentage correct for the program writing measure and Table 13 presents the analysis of variance. A significant test period effect was found, \( F(1,12) = 30.07, p < .0001 \). This indicated a significant drop in performance on this measure.

Tests on proportions of trainees meeting criteria

To test the hypotheses regarding the proportion of trainees exceeding predetermined criteria, the inferential technique for testing differences between proportions for dependent samples suggested by Glass and Stanley (1970) was used. The statistic used in this test is a Z score. In order for the statistic to have an approximately normal distribution with mean = 0 and standard deviation = 1, the sum of the cells being directly tested must be equal to or greater than ten. To fulfill this requirement, the tests were performed by collapsing the groups within each test period and then testing differences amongst measurement periods.
Table 12

Mean Percentage Correct on Program Writing Measure

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>Mid</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>87.25</td>
<td>78.00</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>96.00</td>
<td>61.33</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>88.88</td>
<td>53.62</td>
</tr>
</tbody>
</table>
Table 13

Results of Analysis of Variance on Program Writing Data

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Group</td>
<td>2</td>
<td>376.579</td>
<td>1.37</td>
</tr>
<tr>
<td>Error</td>
<td>12</td>
<td>274.059</td>
<td></td>
</tr>
<tr>
<td><strong>Within</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Period</td>
<td>1</td>
<td>4424.020</td>
<td>30.07*</td>
</tr>
<tr>
<td>Period x Group</td>
<td>2</td>
<td>489.871</td>
<td>3.33</td>
</tr>
<tr>
<td>Error</td>
<td>12</td>
<td>147.122</td>
<td></td>
</tr>
</tbody>
</table>

* P < .0001
**Proportion of trainees exceeding terminology test criterion**

Table 14 presents a series of contingency tables that indicate the frequency of trainees passing or failing the 80% criterion for this test, contingent upon when the test was administered. The test on proportions described above has a special formula which allows the test to be calculated directly from the contingency tables. The results of this test on proportions revealed significant differences between pre-test and mid-test, $Z = 4.12, p < .001$, and between pre-test and post-test, $Z = 4.24, p < .001$. Support was thus found for the hypotheses that the training program would increase the proportion of trainees exceeding the criterion of success from pre-test to mid-test and post-test.

**Proportion of trainees exceeding DTF criterion**

Table 15 is a series of contingency tables that indicate the frequency of trainees passing or failing the 80% criterion for this test, contingent upon when the test was administered. The results of the test on proportions revealed significant differences between pre-test and mid-test, $Z = 3.46, p < .001$, and between pre-test and post-test, $Z = 3.46, p < .001$. These results support the hypothesis that the proportion of trainees exceeding the criterion would be greater at mid-test and post-test than at pre-test.

**Proportion of trainees exceeding program writing criterion**

Table 16 shows the contingency table for trainees at mid-test and post-test who passed or failed the 70% criterion set for this measure. The proposed hypothesis regarding this measure could not be tested due to the lack of pre-test data. However, the maintenance of the results obtained at the completion of the workshop can be examined. The results revealed a significant difference between the mid-test and post-test, $Z = 2.83, p < .005$. This result indicates a significant decrease in the proportion of trainees
Table 14

Contingency Tables of Trainees Meeting 80% Pass Criterion for Terminology Measure

<table>
<thead>
<tr>
<th>Mid-test</th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>0(^a)</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>10</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Post-test</th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>9</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Post-test</th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>4</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mid-test</th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

\(^a\)Number of trainees passing pre-test and failing mid-test
Table 15
Contingency Tables of Trainees Meeting 80% Pass Criterion for DTF Measure

<table>
<thead>
<tr>
<th></th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mid-test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass</td>
<td>$0^a$</td>
<td>6</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fail</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Post-test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fail</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Post-test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Mid-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fail</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

$^a$Number of trainees passing pre-test and failing mid-test
Table 16

Contingency Table of Trainees Meeting 70% Pass Criterion on Program Writing Measure

<table>
<thead>
<tr>
<th></th>
<th>Fail</th>
<th>Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>8(^a)</td>
<td>6</td>
</tr>
<tr>
<td>Mid</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\)Number of trainees passing Mid and failing Post
exceeding the criterion at post-test compared to mid-test.

**Consumer Satisfaction**

Participants used a 5 point rating scale to indicate their satisfaction with the workshop and follow-up visits. All ratings were done anonymously to limit the possible coercive bias inherent in this type of rating. Table 17 shows the mean consumer satisfaction ratings of the workshop and follow-up visits for the four training programs. The range of mean ratings for the workshop was 4.2 to 4.5, and the range for the follow-up visits was 3.7 to 4.7.

Two one-way ANOVA's were computed to test for differences among the groups. Table 18 presents the analysis of variance for both sets of consumer satisfaction ratings. For the workshop ratings, the result was non-significant, $F(3,34) = 1.23$, $p = n.s$. The results for the follow-up ratings were also non-significant, $F(3,24) = 1.89$, $p = n.s$. These consumer ratings results indicate a consistent, high degree of satisfaction with the training program, across all groups at both rating periods.
Table 17
Mean Consumer Satisfaction Ratings for the Workshop and Follow-up Visits

<table>
<thead>
<tr>
<th>Group</th>
<th>Workshop</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.4</td>
<td>4.4</td>
</tr>
<tr>
<td>2</td>
<td>4.5</td>
<td>4.7</td>
</tr>
<tr>
<td>3</td>
<td>4.5</td>
<td>3.8</td>
</tr>
<tr>
<td>4</td>
<td>4.2</td>
<td>3.7</td>
</tr>
</tbody>
</table>
Table 18

Results of Analysis of Variances on Workshop and Follow-up Consumer Satisfaction Data

ANOVA of Consumer Satisfaction Ratings for the Workshop

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Groups</td>
<td>3</td>
<td>0.108</td>
<td>1.23*</td>
</tr>
<tr>
<td>Error</td>
<td>34</td>
<td>0.133</td>
<td></td>
</tr>
</tbody>
</table>

*non-significant

ANOVA of Consumer Satisfaction Ratings for the Follow-ups

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Groups</td>
<td>3</td>
<td>0.747</td>
<td>1.89*</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>0.345</td>
<td></td>
</tr>
</tbody>
</table>

*non-significant
DISCUSSION

The results of this study indicated that the Provincial Team's training program made a significant positive impact on the trainees' skill in using DTF teaching and their knowledge of behavioral terminology by the completion of the workshop. Further, the results suggested that the trainees' knowledge of behavioral terminology and use of DTF teaching were maintained at post-test. As no pre-test measurement was done for the program writing measure, it is not known what impact the training had in this area. All three groups included in the limited analysis on program writing were above the 70% pass criterion set for this measure by the completion of the workshop. However there was a significant decrease in performance on this measure at post-test. The consumer satisfaction ratings of the workshop and follow-up visits were uniformly positive across all groups.

As indicated in Table 7, not all participants who attended the workshop were included as trainees for data analysis. Attrition and staff turnover were the two major reasons for the dropout of direct service staff who did attend the workshop. George and Baumeister (1981) found a very high (73% per year) turnover rate for managers of community based resources for the mentally handicapped. Data were not kept on attrition rates in communities where the training programs were conducted. The turnover rate appeared to be high, although lower than the rate reported by George and Baumeister (1981). For each measure taken separately the rule of inclusion guarded against a bias over test periods due to selective attrition at any one test period (Campbell & Stanley, 1966).
The importance of teaching appropriate terminology is suggested by the widespread use of programmed tests (Becker, 1971; LaVigna & Traphagen, 1978; Patterson, 1975), summary quizzes (Watson, 1973), and use of terminology and short answer questions as units of other behavior modification training programs (Gardner, 1972; Hall, 1978; Schinke & Wong, 1979; Watson, 1972; Weissman-Frisch et al., 1980). As suggested by Watson (1972), the rationale behind encouraging trainees to learn appropriate terminology is that the conceptional learning will allow the trainees to more fully understand the techniques they use.

Increases in the trainees' terminology scores were probably attributable to the combined effects of the team modeling appropriate use of terminology, and the corrective feedback the trainees received concerning their own use of terminology. The differences in mean scores on the terminology test from pre-test to mid-test and post-test were significant. An inspection of Table 8 shows the group means at mid-test and post-test fluctuated around 80%, the criterion for success. Inspecting the data from Table 14 the proportions of trainees achieving the 80% criterion on the test indicates that the majority of trainees passed the 80% criterion at mid-test, regardless of their pre-test scores, and maintained their scores above this criterion from mid to post-test.

A possible explanation for these results is suggested in a study reported by Watson (1972) in which the influence of a reinforcement contingency upon quiz performance was investigated. Progress through the training units was either contingent or non-contingent upon successful completion of the quizzes. Within the non-contingent phases test performance was highly variable, while under the contingent condition the performance was very consistent and above criterion. As the Provincial Team's training program was voluntary, it was similar to the non-contingent
condition. It may be that the variations in the trainees' terminology scores were due to a lack of contingent reinforcement which could override the effects of individual motivational differences.

The most impressive results of this study were those obtained by the measurement of the trainees' DTF teaching skill. These data were in substantial agreement with those presented by Koegel et al. (1977, 1978). The results from the present study were obtained with approximately 7 hours of instruction within the workshop component. Koegel et al. (1977) used approximately 25 hours of instructions. Thus the results from the present study indicated that an acceptable level of skill in DTF teaching can be obtained with much less training than previously indicated.

The post-test data from the present study revealed a maintenance level of over 90% correct in the use of DTF teaching for the majority of trainees. This extends the previous data reported by Koegel et al. (1977, 1978) on the impact of training direct service providers. These two prior studies utilized a post-test immediately following training. The results of the present study suggest that the maintenance of this skill was very good up to the termination visit, 6 to 7 months after initial training.

Two possible factors which would support the maintenance of the skill are suggested. First, as indicated in the Koegel et al. (1977, 1978) studies, the effect of the proper use of this teaching technique is to dramatically increase the child's rate of correct responding. Second, from clinical observations made during this study, there is often a reduction in the student's rate of inappropriate behaviors concomitant with the use of DTF teaching. The immediately reinforcing value of these
two factors for the trainee would support a high maintenance level.

The results from the program writing mid-test indicated a high level of trainee proficiency in this skill, although the post-test scores revealed a significant decrease in proficiency. As indicated in the results, only three trainees completed the pre-test measures. The assumption underlying the gathering of the program writing data was that the trainees would be writing programs which could then be scored with the program writing measure. This assumption was based on the piloting of the program writing measure prior to this study. Why the assumption was justified with the pilot sample but not the experimental sample can be explained as a function of a history bias. Many of the pilot trainees had been exposed to or trained in program writing prior to testing. For instance, several pilot trainees had received training from Laurel House, a behaviorally oriented treatment center which uses program writing as an essential tool for structuring their training of the mentally handicapped. Very few of the experimental trainees had any such history. Thus, a program writing measure, which could be used generally to gather pre-test data, would need to assume that most trainees would not be actively writing programs.

An assumption which could be made about this pre-test situation is that trainees' pre-test scores on a program writing measure were low. A recently reformulated program writing measure now assumes that trainees will not be writing programs. The trainees are now given the structured task of writing a program during the site visit. Pre-test data using this new measure have been collected on two training programs by the Woodlands Team. The mean pre-test score was 18.5% with a range of 6% - 58%. As the definitions of trainees included for evaluation were the same for the Woodlands and Provincial Teams, this data offers some post hoc support for
the assumption that the trainees for the present study would also have had low pre-test scores.

Several factors appeared as possible agents contributing to the clinically and statistically significant reduction in trainees' scores following training. As program writing was a new and difficult skill for most trainees, continuous reinforcement would be necessary to encourage its continuation. There appeared to be few immediate rewards for program writing. The reinforcement of the skill would have to be managed and programmed until the natural nonprogrammed reinforcers for this skill could be established. However, in the normal work setting of the trainees, there seemed to be almost no reinforcement and sometimes punishment, for this skill. For instance, the trainees working on a contractual basis in several instances had contracts which precluded payment for any activity not related to working directly with the client. For teachers, at the beginning of each school year, commitments were made to classroom and school routines which made the changes to an individualized program basis an onerous task.

Testing was the major threat to the internal validity of this study. The present study used a repeated measures design, with measures having reactive properties. Multiple exposures to the measurement devices could have had a biasing effect on the data. It is not possible to rule out this threat from the terminology and DTF measure as the generally consistent rapid increase in mean scores at mid-test which were maintained at high levels at post-test could potentially be explained as a result of a testing bias. The significant decrease from mid-test to post-test on the program writing measure could also be explained as a result of a testing bias.

Threats to the external validity of this study come from a possible selection x training program interaction and the reactive effects of testing. Two types of selection bias are apparent in this study which
limit the generalizability of the results. All trainees were volunteers for this training program and only data for trainees who completed all three test periods on a dependent measure were included in this study. With these two biases, generalizations to non-volunteer trainees and trainees who do not complete the entire training program would seem inappropriate. The reactive nature of the measures used in this study made it uncertain whether the data could be generalized to training conditions where no measurement was taken or unobtrusive measurements were used.

The present study extends the previous literature on training programs for direct service workers of severely mentally handicapped clients. This study, with its cyclical delivery of a training program, indicated that an effective training program, based on methods suggested by previous research, can be delivered to a wide variety of trainees in different settings. The major area of weakness was the maintenance of the trainee's program writing skills. The portability of the training program, where three trainers could train 30-40 trainees a year in their home communities, would suggest that this was a very cost-effective method for delivering a training program. The cost-effectiveness of such a training program would be particularly evident in an area with a widely dispersed population, such as is the case in the Province of British Columbia.

The use of DTF teaching as the major teaching technique focused on during the training program extends the research previously done on the technique (Koegel et al., 1977, 1978). As mentioned earlier, initial high levels of proficiency were obtained in the present study using substantially less training time than Koegel et al. (1977). Also, it was demonstrated that the high level of proficiency was maintained for
up to 6 or 7 months following the workshop. The previous studies by Koegel et al. (1977, 1978) had not studied the long term maintenance of this skill.

Three areas can be identified for further research. Although prior research (Fredericks et al., 1979; Gladstone & Sherman, 1975; Koegel et al., 1977, 1978; Panyon & Patterson, 1974; Schinke & Wong, 1977) has shown that training direct service staff has a beneficial impact on the target clients, this has not been demonstrated for the present training program. Until such time as the Provincial Team's training program can be demonstrated to lead to beneficial changes for the target clients, its value for the target clients is uncertain. A more appropriate measure of program writing skill which is capable of reflecting the present level of this skill in the trainee population is required. Following the development of such an instrument, research on methods for maintaining this skill would be of interest. Finally, the training program of the Provincial Team was a complex package. The use of a dismantling treatment strategy (Kazdin, 1980) to investigate the possibility of reducing the complexity of the package would have both theoretical and practical benefits.
REFERENCE NOTES


REFERENCES


Panyon, M., & Patterson, E. Teaching attendants the applied aspects of behavior modification. *Mental Retardation*, 1974, 12, 30-32.


APPENDICES
Manual for PIRT

Training Program

Douglas Lee
1980
Site Visit

The purpose of the site visit is to finalize the organizational and administrative requirements in preparation for the workshop. The team member acting as contact person is responsible for meeting with the intake and review committee to make sure that all the necessary arrangements are done for the workshop. (e.g., a room is booked, babysitters and childcare workers have been found, and substitute replacements for trainees are available). As the specific requirements of each community varies widely, it is usually necessary to draw up an unique list of requirements which must be fulfilled for the workshop.

All team members are responsible for a certain portion of the trainees and target clients during the site visit. For each trainee, 8 items must be covered:

1. Behavior terminology test administered.
2. DTF teaching video (10 consecutive trial minimum completed)
3. Copy of the trainees most recent written program collected.
4. Video release is signed.
5. Academic record sheet completed.
7. Program text given (money collected if not a MHR employee).
8. Trainee information sheet completed.

All the items except the video taping are completed during an interview that is either one to one or in a small group (e.g., 3 trainees from the same work setting may be seen together).

As the requirement for the site visit will have been specified in some detail during the pre-referral visits, it is rarely necessary to give
more than the briefest overview rationale for requesting the information at the site visit. However, as many trainees have not been videotaped before or seem anxious about being videotaped, as much time as required is spent in making sure the trainees are as relaxed as possible. One way of dealing with the trainee's anxiety is to reassure her or him that the purpose of the videotaping is to help the team gather information which will be used to better tailor the content of the workshop to the trainee's own needs, and to turn on the camera and then walk away from the camera. Also, as the videotaping is being done with the trainee working with a target client, it often appears that the trainee's perception is that the video is more on the target client than on herself.

The issue of the trainee's perception of the team's focus is an area in which the team often times must spend considerable energy. It has become very obvious that the common perception of consultants coming into a community is that their focus will be almost exclusively on the target client. The consultant will come in, "assess" the client, write out a diagnostic/prescriptive report, and then leave. The team's perspective is that they will deal with the client as much as possible only through the trainee. It is the trainee who is the direct target of the team with the target client being the indirect target. Helping the trainees' make this shift is very important in order that their expectations be supportive of the events to take place during the workshop. In previous training programs when much less effort was directed towards ensuring the trainees' perceived themselves as the direct target of the team, some trainees would express dissatisfaction with the team for spending so much time with them and so little time working with target clients.
The other major responsibility of the team is to observe the target client. Prior to the site visit, as much written assessment material as possible has been gathered on the target client. These reports are helpful in providing general information. The observations by the team member are aimed at gathering present functional information on how the target client deals with his present environment. Some general areas of interest are:

1) how does the client interact with the teachers/trainers
   a) during instruction
   b) during non-instruction
2) how does the client interact with other people
3) how does the client interact with the physical environment
   a) task materials
   b) non-task materials

Due to the idiosyncratic nature of this particular target client populations' learning history, learning strategies, and perceptual abilities, just what areas will be focussed on is impossible to detail a priori.

During or following this observation period, it may arise that clarification of what is being observed can only be done by actual interaction with the target client. Often this is done to assess how the client will react to different teaching strategies. As the trainees usually do not have a variety of different teaching strategies in their repertoire the team member must do the assessment directly with the target client. Occasions have arisen in the past where the more important question is "What will happen to the trainee/target client interaction if it is the trainee who actually implements the new procedure?" An example
of where this may be the question is when the interaction is such that
the trainee is getting the exact opposite response they want and this
appears to be due to the positive and negative contingencies being
misapplied to the target client's behavior. Since the two people in
this interaction may have been involved in this S - R - C system for
an extended period of time, it is possible that the interaction is
perceived as a formalized structure by the client. The effects of a
new person (team member) altering this structure may be totally different
than having the trainee alter the structure of the interactions.

In order to accomplish the above mentioned items, the team members
usually book approximately 1-½ hours per trainee. This is broken into
introduction (outlining purpose of visit and rapport building),
observation and videotaping, and meeting to complete various forms, and
the terminology test. As the team is usually at a site for 2-3 days, it
is possible that more observation, videotaping or discussion with the
trainees will be scheduled during the visit.

Workshop

Due to the variability between training sites on such factors as
the stated needs by the trainee, functional level of target clients,
language capability of target client, size of training group, and the role
the trainees have within the service delivery context for the target clients,
the specific content of a workshop is quite variable. However, there is
a structure to the workshop which is identifiable over and above this
contextual variability. It is important to remember that the structure
and the content to be presented will be somewhat idealized in order to
present as clear a model of the workshop format as possible. The description
will be broken down into the five days of the workshop.
Day 1

9:00 - 10:15  Introductions:  As the trainees may be from several community facilities and the team members may not have met all trainees during the site visit, quite formal introductions are necessary in order to start developing a sense of group cohesion and as an initial rapport step.

Goals: From the trainee's expressed needs, the observational data from the site visit, and review of the assessment materials on the clients the team will draw up a series of goal statements which attempt to pull together the information from these sources. Since the workshop is not the end of the training program, but rather the intensive initial learning phase of a protracted training involvement between the team and the trainees it is also necessary to state along with the goals of the workshop, how the workshop fits into the entire training program. An essential feature to point out is that the trainees are not expected to be able to immediately and independently utilize everything that will be taught during the workshop. It has happened with some frequency that trainees come to view the workshop as the end of the team's involvement which leads to an overwhelming sense that they cannot possibly learn "all this" so fast. This sense of being overwhelmed can greatly interfere with the person's acquisition of the material. Often it is necessary to reiterate for individual trainees during the workshop that the team will be back several times over the next few months.

Autism Overview: Although most of the trainees will be working with clients who have autistic or autistic-like behaviors, it is quite common that the trainees will not have had a descriptive
overview of the syndrome. A film called "Bobby" and/or a verbal overview presentation on autism and the behavioral therapy techniques used with autistic people are given. Even if several trainees in a workshop have an adequate knowledge of this area, the overview is still done to try and provide some uniformity to the trainee's level of information and also to indicate the team's behavioral perspective.

10:15 - 10:30 COFFEE

10:30 - 12:00 Introduction to DTF teaching: There are several components to this introduction.

1) "Denise" VTR - a training tape made by PIRT and WIRT which gives a rationale for this type of teaching, defines the components of the technique and gives examples of its use.

2) The team will demonstrate the use of DTF by 2 members role-playing a teaching situation. The behavior which is used in this role-play will be one which has been mentioned by the trainees as a behavior they have trouble teaching. This demo is followed by a discussion.

3) A further demonstration will be done. The trainees will be broken down into small groups with a team member and the trainees will take turns role-playing teaching the behavior from the previous demonstration. This will be followed by a discussion.

4) As the majority of trainees will be experiencing difficulties with their clients because they have not found a method for training the client in the prerequisite behavior for working on a task (e.g. attention to task, sitting still for several minutes), a small lecture will be given on what prerequisite behaviors are
and how they relate to the teaching task. As the prior demonstrations or roleplay is often centered on teaching eye-contact, sitting in a seat or keeping your hands still, the lecture is often an elaboration of the prior exercises.

12:00 - 1:15  LUNCH

1:15 - 2:30  Following-up from the mini-lecture on prerequisite behaviors, the team will demonstrate by role-play, the teaching of eye-contact or sitting still. This is followed by the trainees breaking into small groups and role-playing the teaching of one of these behaviors.

2:30 - 2:45  COFFEE

2:45 - 3:15  Up to this point the focus has been upon the face-to-face teaching interaction. At this point, the team will role-play a teaching session, introducing the concept of pass and fail criteria. After a brief introduction to the concept, the team members will do a role-play in which the "student" will pass from one step to the next and will fail one step, necessitating a return to a previous step and a review.

3:15 - 3:45  A film, "Teaching Makes a Difference #1" is shown to summarize much of the day's work.

3:45 - 4:00  Homework assignments: This will involve a review of chapters in the Teaching Makes a Difference manual in preparation for Day 2, plus 1 or 2 other relevant readings.

Day 2

9:00 - 9:30  The day opens with a discussion period to clarify questions from the previous day. If a definition list was given out, this is reviewed at this time.
9:30 - 9:45  A mini-lecture which defines what a baseline is and a rationale for using a baseline is given.

9:45 - 10:15  The trainees break into small groups with a team member. The target clients are introduced into the workshop environment at this point to start the direct practice sessions between the clients and trainees. This practice session will focus upon the collection of baseline data usually on an attending/prerequisite behavior. Each trainee will do 10 trials with the client. This is the first session in which the trainee and clients will be videotaped during the workshop. All other direct practice sessions will also be videotaped. The videotaping will be rotated through the groups. Not all trainees will be videotaped each session. The baseline data collected on the target clients will be used as the baseline data for the workshop and it is stressed that if the trainees want to work on this target behavior after completion of the workshop, a new baseline will have to be done.

10:15 - 10:30  COFFEE

10:30 - 11:00  A lecture is given introducing program writing. Trainees are given an example program and the program writing guide. Using an enlarged version of the example program, each component of the program is discussed. The format of the program writing sheet is very similar to that presented in the "Teaching Makes a Difference" manual. Particular stress is laid on the "how to" of writing a good program.

11:00 - 12:00  The trainees break into small groups and start to write a program. The intent of this session is not to complete an entire program, but to have enough of the program outlined so that it can be implemented with the clients after lunch.
12:00 - 1:15 LUNCH

1:15 - 2:00 Direct practice session with clients using program outlines written before lunch. This is a particularly useful point for highlighting confusions and ambiguities in the program that is being used as the trainees often have used unclear wording or ill-defined terms in their program. However, once other people are attempting to utilize their programs the need for explicit statements becomes evident. Throughout the directed practice sessions occasions arise which enable you to compare the trainee's expectations of the client's ability to perform a certain task (by what they write in their programs) and the client's actual ability level. The benefit of having explicit expectations in a program (e.g., how much prompting is required, pass/fail criteria) for comparison to the client's actual behavior can be exemplified. This could take the form of suggesting to the trainee that the program (especially a first draft) is your best guess of how the client will learn this task. How well the client does (as noted from your data) is feedback as to your best guess. By altering components of your program until you get the data you want, two results will occur: the client will learn the behavior and you will end up with a written record of an effective teaching strategy.

2:00 - 2:30 Film "Teaching Makes a Difference #2". This film focuses on the issues of how to build a programmed curriculum for each individual into a special classroom environment.

2:30 - 2:45 COFFEE

2:45 - 3:45 A lecture/discussion is done focussing on maintenance and generalization issues, and, rewards and reinforcement schedules. All these issues are dealt with together in order to provide a guide
of how the trainees can move their clients from the schedule of reinforcement and types of rewards used to teach a behavior into schedules and rewards which will maximize the maintenance and generalization of the new behavior.

3:45 - 4:00 Homework assigned: review chapters in manual for tomorrow, complete definitions list; either complete the program started today, or write one for a new behavior for a different client.

Day 3

9:00 - 9:15 Hand in programs written as homework. Review definition list with discussion.

9:15 - 10:00 The program components of prompt, probes and branching were outlined in the program writing lecture on Day 2. They are now gone into in more detail. Specifically various different prompts are outlined (manual, proximity, and within stimulus) and more detailed rules for probing and branching are given.

10:00 - 10:15 During the previous hour, two team members would have reviewed the programs, chosen the best ones for the clients present today, scored them, and written comments. These programs are presented to the group.

10:15 - 10:30 COFFEE

10:30 - 11:45 The trainees break into small groups, each group with a program. The trainees first role-play the program to become familiar with it and then have a directed practice session with the client, attempting to follow the written program.

11:45 - 12:00 During the preceding practice session, one of the team members has finished marking all the programs. The programs with
written feedback are given back to the trainees. During this period the marker reviews the programs in general for the whole group. Questions on specific questions will be dealt with throughout the day.

12:00 - 1:15 LUNCH

1:15 - 2:00 Lecture on language/communications training. Due to the highly variable language and communication abilities of the clients, this lecture will often change drastically from workshop to workshop. However, a major intent of this lecture is to focus on the need for developing a communication method which each client can use. As many clients will often exhibit almost no structured communicative ability, time is also spent on how to start teaching language to someone with little or no skills in this area.

2:00 - 2:30 Either a role-play demonstration or a video-tape presentation will be done to exemplify teaching strategy issues related to communication training.

2:30 - 2:45 COFFEE

2:45 - 3:45 A direct practice session with the clients focussing on teaching communication behavior. This practice session is used for gathering ideas to write a language program for homework.

3:45 - 4:00 Homework; review chapters in manual. Write a language program. Complete definition list.

Day 4

9:00 - 9:15 Hand in homework, review definitions, and discussion.

9:15 - 10:00 The principles of behavior management are outlined in a lecture format. As there is a session later in the day focussing on techniques of behavior management, this present session is more
on defining operant principles than techniques.
10:00 - 10:15 Discussion of scored programs to be used for directed practice.
10:15 - 10:30 COFFEE
10:30 - 11:45 Role-playing programs and then practicing the implementation of the program with the clients.
11:45 - 12:00 Programs done for homework handed back with general comments to the trainees.
12:00 - 1:15 LUNCH
1:15 - 2:15 Techniques of Behavior Management lecture. Source material can be found in LaVigna & Donnellan (1976). As a focal point of the talk, an example behavior management program is given to the trainees. An oversized version of the example is used by the team member for reference purposes. An explicit bias of this session is on non-aversive (particularly avoiding intrusive methods) methods of behavior management. As a behavioral observation exercise, videotaped segments are used, where the trainees are asked to count the frequency of a certain behavior during several 2 minute segments. Each person's data is reported which allows for a discussion of observer reliability and to highlight that behavior naturally occurs at a variable rate.
2:15 - 2:30 COFFEE
2:30 - 3:00 An introduction to task analysis is given with its relationship to program writing. A task analysis exercise is done where the trainees break into dyads and with their backs turned to each other, each trainee, in turn, must describe step by step how
the other trainee is to do a task.

3:00 - 3:45 Film "Teaching Makes a Difference #3". This film overviews many of the areas covered in the previous two days.

3:45 - 4:00 Homework. Review chapters in Manual. Do a task analysis or write another program. Complete definition list.

Day 5

9:00 - 9:15 Hand in homework, quiz and discussion on definition list and readings.

9:15 - 10:15 Curriculum lecture. Over the past four days, the components which are required to make a curriculum for an individual client have been taught. These components include such content areas as prerequisite behaviors, language skills, behavior management and self-help skills. How to write programs, collect data, set criteria, use shaping, DTF, and schedules of reinforcement are some of the techniques focused on during the workshop. The intent of this lecture is to provide the trainees a schema for pulling all this information together to make an overall plan for each client. Much of the material covered is from the literature on program writing. Since putting an individual curriculum together is a very complex task, this session is very much a conceptual introduction into curriculum. Many of the pragmatic issues of curriculum construction are dealt with during the follow-up visits.

10:15 - 10:30 COFFEE

10:30 - 11:45 Small group planning is done at this time in preparation for the first follow-up. The intent of this planning is to set out what each trainee is going to attempt between the workshop and the
first follow-up. Specific requests for information are noted
so the required materials can be ready for the follow-up visit.
As it is obvious that the trainees could not hope to use all the
information and skills learned during the workshop immediately,
an attempt is made to have the trainees focus on one client,
one behavior across 2 or 3 clients, or on using one teaching
technique. The integration of the other information from the
workshop will be done in a step wise fashion at a pace that is not
too disruptive to the trainee's work setting.
11:45 - 12:00 Homework feedback. General feedback is given regarding
the task analysis or programs done as homework.
12:00 - 1:15 LUNCH
1:15 - 1:45 Lecture on token economy. After defining what a token
economy is, the time is spent on how to build one. This goes from
the initial pairing of the token to a primary reinforcer to full token
economics with several reinforcers costing differential amounts.
1:45 - 2:30 Each evening the videotapes of that day's direct practice
sessions were scored. From the previous 3 days of videotapes, general
comments are pulled together to give to the group as a whole. Individual
feedback is also given. If it is possible to find select segments
of videotape that highlights the comments made, they are shown after
requesting the trainee's permission.
2:30 - 2:45 COFFEE
2:45 - 3:00 The terminology mid-test is given.
3:00 - 3:45 The written trainee feedback is done at this time, followed
by an open feedback and comments period with the whole group.
3:45 - 4:00  A wind-up period used for closing off the workshop and for final preparations for the follow-up visit.

**Follow-up Visits**

The structure of the follow-up visit is more variable than that of the workshop. A main factor for this variability is that the focus during the follow-ups is towards helping each individual trainee generalize the skills and information of the workshop to their normal work setting. As noted in the small group planning session on Day 5 of the workshop, each trainee has specific goals set for themselves.

At an abstracted level, the flow of the follow-up could be characterized as moving from the specific to the general. For example, the goals set for the trainee for the first follow-up are specific to one client, one type of behavior or one teaching technique. The intent is to move from this starting point to other clients, types of behavior, or teaching techniques. The rate of movement however, will be highly variable due to the Environment x Client x Trainee interaction.

Recognizing the diversity of the follow-up goals, there are several common themes followed during the large majority of follow-up visits. These are:

1) Direct observation of the trainee working with a client. As the trainee is usually requesting a team member to observe a particular teaching session because something isn't working, the team member may role-play or model with the client what it is that could be done differently. This observation will also include checking on whether there is a written program for this behavior and reviewing the data that has been collected since the previous visit.
2) Program Writing: has the trainee written any programs since the previous visit? If so, give feedback. If not, what has stopped the trainee from doing so. Do the trainees want any input on altering an existing program. The programs written by trainees during follow-up are often sketchier than those done in the workshop. This usually gives ample opportunity for discussing many relevant issues. Common components left out are generalization, maintenance, pass-fail criteria and branching or alternate steps if the program does not work. These issues often can lead into discussions of long term goals (curriculum) and issues pertaining to how the trainee perceives the client's learning style.

3) Data: although the collection of data is part of the teaching and program writing techniques taught during the workshop, special emphasis is placed on data during the follow-up visits. There are two reasons for this. First, data collection is very foreign for most trainees and without support they could not collect any data. Secondly, data is seen by the team members as a cornerstone and feedback mechanism for so much of what they wish to accomplish with the trainee.

4) Curriculum: The issues surrounding curriculum often develop over the entire series of follow-up visits. If curriculum is to be a major issue for the trainees, it usually becomes important during the third or fourth visit. Such package curriculums as the Behavioral Characteristics Progression (BCP) and Student Progress Report (SPR) are often brought along on a follow-up visit as an aid in helping the trainee develop a structured approach to building an individualized curriculum.
5) Behavior Management: A troublesome issue for many trainees is how to deal with the behavioral excesses or deficits which are often exhibited by this client population. It is a particularly thorny issue in that, in the long run, the best method for dealing with these behaviors is to put the client in an enriched productive training environment suited to his specific needs. In the short term, however, the trainee has an almost impossible time implementing the teaching strategies which could lead to a positive learning environment because of the behavioral excesses or deficits. It is very important that the work and energy which is taken on dealing with a behavioral excess or deficit not overshadow the work needed to build a positive teaching environment.

In carrying out a follow-up visit, it is important to reinforce the many good things the trainees are doing. Without the support of the team members the trainees can quickly become overwhelmed by the task of integrating all the materials presented by the team into everyday practice. Also, because so many of the techniques taught to the trainees are made up of several simple components, the trainees can easily lose track of how much they have learned since each component is so easy to learn. It often takes the perspective of an outsider to help the trainees appreciate how much they have learned.

Because there will only be five follow-up visits, the team must start early on reinforcing independence by the trainee. This independence is particularly related to the trainee initiating requests rather than waiting for a suggestion from the team. Once the termination visit is completed, any further contact between the team and the trainees will only occur if the trainee initiates the request.
Termination Visit

Besides being the conclusion of the training program, this visit is the time for gathering the post-test data. A copy of the most recent program written by the trainee is collected, the terminology test is readministered and a videotape segment is done of each trainee's teaching (a minimum of 10 trials). As well, the trainees are asked to complete a follow-up feedback form.

Since the follow-up visits have been scheduled at progressively longer periods with the termination visit being approximately 8 weeks after the last follow-up, there will have already been a naturally occurring distancing between the trainees and team members. The conclusion of the training program is then not a sudden finish, but rather a cumulative end point. Every effort is made at this point to indicate that the team's role is now changing to a consultive role and that further contact will occur only at the request of the trainee.
APPENDIX B
Provincial In-Service Resource Team

NANAIMO WORKSHOP

November 26th - 30th 1979

MONDAY, NOVEMBER 26TH 1979

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00</td>
<td>Introduction and Schedule</td>
</tr>
<tr>
<td>9:30</td>
<td>&quot;Bobby&quot; Film</td>
</tr>
<tr>
<td>9:45</td>
<td>&quot;Why D.T.F.&quot; Autism Therapy and Teaching Techniques</td>
</tr>
<tr>
<td>10:15</td>
<td>Coffee</td>
</tr>
<tr>
<td>10:30</td>
<td>&quot;Denise&quot; VTR and &quot;Group Teaching&quot; VTR utilizing DTF components</td>
</tr>
<tr>
<td>10:45</td>
<td>D.T.F. Demonstration</td>
</tr>
<tr>
<td>11:00</td>
<td>Discussion of Demonstration</td>
</tr>
<tr>
<td>11:10</td>
<td>D.T.F. Demonstration</td>
</tr>
<tr>
<td>11:20</td>
<td>Mini Lecture Teaching Pre-requisite Skills</td>
</tr>
<tr>
<td>11:50</td>
<td>Discussion</td>
</tr>
<tr>
<td>12:00</td>
<td>Lunch</td>
</tr>
<tr>
<td>1:15</td>
<td>Demonstration of Teaching Attending Skills</td>
</tr>
<tr>
<td>1:45</td>
<td>D.T.F. Role Play eye contact Programme</td>
</tr>
<tr>
<td>2:30</td>
<td>Coffee</td>
</tr>
<tr>
<td>2:45</td>
<td>D.T.F. Role Play with Criteria</td>
</tr>
<tr>
<td>3:15</td>
<td>T.M.D. #1 - film</td>
</tr>
<tr>
<td>3:45</td>
<td>Homework Assignment: Review Chapters 2 &amp; 4 &amp; 6 in manual</td>
</tr>
<tr>
<td></td>
<td>Read Kozloff Handout</td>
</tr>
<tr>
<td></td>
<td>Complete Definition List</td>
</tr>
</tbody>
</table>
NANAIMO WORKSHOP

TUESDAY, NOVEMBER 27th 1979

9:00 - 9:30 Quiz and Discussion
9:30 - 9:45 Mini Lectures - Baselining & Assessment
9:45 - 10:15 D.T.F. Hands on - Baselining Eye Contact
10:15 - 10:30 Coffee
10:30 - 11:00 Programme Writing Lecture
11:00 - 12:00 Programme Writing
12:00 - 1:15 Lunch
1:15 - 2:00 D.T.F. - Hands on with Programmes
2:00 - 2:30 T.M.D. #2
2:30 - 2:45 Coffee
2:45 - 3:30 Generalization, Rewards & Reinforcement Schedules, and Maintenance
3:30 - 3:45 Discussion
3:45 - 4:00 Homework: Review Chapt. 5, 12 in manual
   Read Reward Handout
   Complete Definitions List
   Finish Programme started earlier including Generalization
HAND IN HOMEWORK
9:00 - 9:15 Quiz and Discussion
9:15 - 10:00 Prompts, Probes, Branching
10:00 - 10:15 Discussion of Programmes
10:15 - 10:30 Coffee
10:30 - 10:45 D.T.F. - Role Playing Programmes
10:45 - 11:45 D.T.F. - Hands on with Programmes
11:45 - 12:00 Homework Feedback
12:00 - 1:15 Lunch
1:15 - 2:00 Teaching Language Lecture
2:00 - 2:15 D.T.F. - Demonstration of Language Programme, V.T.R.
2:15 - 2:30 D.T.F. - Role Playing with Language Programme
2:30 - 2:45 Coffee
2:45 - 3:45 D.T.F. - Hands on re: Language, Assessment, Teaching and Programme Writing
3:45 - 4:00 Homework: Review Chapter 3
Write Language Programme
Complete Definition List
NANAIMO WORKSHOP

THURSDAY, NOVEMBER 29th 1979

9:00 - 9:15  Hand in Homework
            Quiz and Discussion

9:15 - 10:00  Principles of Behavior Management

10:00 - 10:15  Discussion of 3 Programmes

10:15 - 10:30  Coffee

10:30 - 10:45  D.T.F. - Role Playing with Programmes

10:45 - 11:45  D.T.F. - Hands on with Programmes

11:45 - 12:00  Homework Feedback

12:00 - 1:15  Lunch

1:15 - 2:15  Techniques of Behavior Management

2:15 - 2:30  Coffee

2:30 - 3:00  Task Analysis

3:00 - 3:45  T.M.D. #3

3:45 - 4:00  Homework:  Review Chapter 8, 9, & 10 - manual
              Do a Task Analysis
              Complete Definitions List
NANAIMO WORKSHOP

FRIDAY, NOVEMBER 30th 1979

9:00 - 9:15  Hand in Homework
            Quiz
9:15 - 10:15 Token Economy
10:15 - 10:30 Coffee
10:30 - 11:45 Small Group Planning
11:45 - 12:00 Homework Feedback
12:00 - 1:15 Lunch
1:15 - 1:45  Trainee's requests for additional information
1:45 - 2:30  Video Feedback
2:30 - 2:45  Coffee
2:45 - 3:00  Mid-test
3:00 - 3:45  Feedback to P.I.R.T.
3:45 - 4:00  Wind up, Preparation for follow-up
For questions with blanks, please place your answer in the blank. For all other questions please circle your choices. This test will take approximately 15 minutes.

1. The teaching procedure where you reinforce closer approximations to the goal behavior is called: -
   a. chaining
   b. shaping
   c. prompting

2. If a teacher or parent does not have complete control over all the reinforcement which maintains an undesirable behavior he/she cannot use ______________ as a procedure for eliminating that behavior.

3. Which of the following is most important in early language teaching: -
   a. teaching a child to label objects and activities in his environment.
   b. teaching a child to distinguish vowel and consonant sounds.
   c. teaching a child to express her/his needs and desires.

4. The ______________ indicates to the student that it is time for a particular response to occur.

5. ______________ is needed in a training program so that we can accurately and reliably assess whether the program is helping the child to learn.
6. It is important to remember that a good assumption to take when teaching an autistic like child is that, she/he will ___ learn what you teach them:
   a. only
   b. eventually
   c. never
   d. always

7. A child's language, which is unintelligible to anyone except someone who knows the child very well, (e.g. the sound "d" to mean down) can be:
   a. taught to other language-delayed children in the class
   b. shaped into more intelligible forms of communication
   c. extinguished by ignoring it
   d. rewarded differentially each time it occurs

8. If you decide upon a set of ____________ for each step of a program, you will more easily be able to decide when to move on to the next step, or when to stop a program.

9. If a behavior (such as self-stimulation) occurs very frequently during the day, _______ is the most appropriate way to collect data.
   a. event recording
   b. time sampling recording
   c. duration recording
   d. anecdotal recording

10. A prompt is used to ____________ a correct response.
11. For many programs, particularly self-help skills _________ chaining is most effective.
   a. backward
   b. successive
   c. simple
   d. contingent

12. Three of the most important pre-learning behaviors which the child must have are: - (circle three alternatives)
   a. can talk or sign with the instructor
   b. can work at some task
   c. has eye contact
   d. does not have any self stimulation
   e. has good fine motor skills
   f. is quiet for a minimum of 30 seconds
   g. cooperates with simple commands

13. The two most important criteria that should be met when using prompts are that they must be _________ and they must be able to be _________.

14. The consequence should be delivered _________ following the child's response.

15. It is essential to get the child's _________ before giving the instruction.
P.I.R.T.

SCORING KEY FOR BEHAVIORAL TERMINOLOGY TEST

1. B
2. Extinction (no other)
3. C
4. S^d; instructional stimulus; instruction; cue
5. Record keeping; data
6. A
7. B
8. Criteria
9. B
10. Insure; assure; elicit
11. A
12. B; C; G
13. Effective: faded; eliminated; extinguished; phased out
14. Immediately; promptly; directly
15. Attention; eye contact
EVALUATION FORM FOR ASSESSING NAME OF PROGRAMME WRITER

DTF PROGRAMME WRITING COMPETENCY

<table>
<thead>
<tr>
<th>TOTAL SCORES</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
</tr>
</thead>
</table>

FACE SHEET

Target Behavior
a. Stated in observable behavioral terms (duration, topography, intensity).
b. Statement of behavior's relation to S or instructional stimulus.

Prerequisite Skills
a. Lists all critical, already learned skills, needed to learn new behavior.

Setting/Time
a. All settings listed as concretely as required for clarity.
b. Times of programme use listed.
c. Any specified settings or times when programme not used are stated.

Materials
a. List includes all required materials.
b. Quantities of all materials given.
c. Important specifications of materials listed.

Baseline
a. Number of trials for baseline.
b. Pass criteria given.
c. Fail criteria given.
d. Results given when possible. If not possible count as N/A (i.e. during workshop).

Probe Procedure
a. All probe steps listed, or the system of probing described.
b. Pass criteria given. are optional if listed in program steps.
c. Fail criteria given.
**PROGRAMME SHEET**

**Follow-up/Generalization**

a. Procedure of monitoring maintenance of behavior.

b. Maintenance pass criteria given.

c. Maintenance fail criteria given.

d. Branch or recycle step given for fail criteria. (Optional, if not stated count as N/A).

e. Outline steps for generalizing programme.

**Instructional Stimulus**

a. Simple – appropriate language level for the child. If is situational, count as N/A.

b. Consistent with goal behavior.

**Prompt**

a. Prompts described in behavioral terms.

b. Prompts faded in consistent manner if not for child’s functioning level. If not required for.

c. Specification of timing of prompt program count (i.e. during forward chaining) a, b, c, as N/A.

**Response**

a. Stated in behavioral, measurable terms.

b. Time limit for response completion given.

c. If time criterion changes that it is done consistently.

d. Topography or intensity of behavior changes consistently towards target behavior.

**Consequence/Schedule**

a. Positive reinforcement specified for correct behavior.

b. Punishment is specified for incorrect behavior.

c. Schedule of reinforcement for correct behavior given.

d. Schedule of punishment for incorrect behavior given. (optional)

**Criterion**

a. Pass criterion for each step given.

b. Fail criterion for each step given.

c. Branching or recycling step/system indicated for fail criteria. (optional).

d. Pass criterion for probes given. Parts d and e are optional if listed.

e. Fail criterion for probes given. on face page.
**PROGRAMME SHEET**

**Data**
- a. Indicate general method.
- b. Indicate specific notation.

**Others**
- a. Probes are specifically notated in program steps (optional if listed on face page)

be sure to add the scores from all pages.

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL ITEMS SCORED

MINUS

TOTAL ITEMS INCORRECT

EQUALS

TOTAL ITEMS CORRECT
EVALUATION OF PROGRAMME WRITING COMPETENCY FOR

D.T.F. PROGRAMME

SCORING KEY DEFINITIONS

Face Sheet

Target Behavior -

a. Stated in observable behavioral terms, with reference to duration, topography, and intensity, if applicable. The statement of duration should include "time to completion" upon the instruction, i.e. Upon the instruction "get dressed", John will be dressed within 5 minutes.

b. Within the description of the target behavior, there is a description of the instructional stimulus or discriminative stimulus for the behavior. (see above example)

Prerequisite Skills -

a. Lists all the critical prerequisite skills the child needs in order to complete the task. These skills are skills the child has already mastered and should not be included in the programme as behaviors to learn.

Setting/Time -

a. All settings for the programme listed as concretely as required for clarity.

b. Times of programme implementation stated i.e. 9:00 - 9:30 a.m.

c. Any specified settings or times when the programme is not to be used are listed i.e. dressing programme not used during swim period.

controlling voice level programme not used during recess.
Materials -

a. List includes all required materials.

b. Quantities of all materials given.

c. Any important specification of materials is listed - i.e.:

   Montessori stacking blocks, not "blocks"

   ski jacket with large zipper, not "jacket"

Baseline Procedure -

a. Number of trials for baseline.

b. Pass criteria given.

c. Fail criteria given.

d. Results given, where possible. If not possible (i.e. workshop programme) count as N/A. If programme is in use but no data given count as error.

Probe Procedure - (optional) If listed in programme steps count as N/A. Mandatory if probe not listed in programme steps.

a. All probe steps are listed, or the system of probing is described, i.e. after completion of any step after step 2 probe for mastery step.

b. Pass criteria given.

c. Fail criteria given.

Follow-up/Generalization -

a. Follow-up procedure described i.e. How to monitor maintenance of behavior after the last step of the programme is completed.

b. Pass criteria given. (Pass criteria means termination of follow-up and programme).

c. Fail criteria given.

d. Fail criteria must be coupled with a branch or recycle step to achieve
mastery level criteria. (Optional). If not stated, count as N/A.

e. The steps for generalizing the programme is outlined, i.e.:
   - changes in reinforcement schedule
   - new related tasks to learn
   - other behaviors to be chained with this behavior.

Programme Sheet

NOTE: If programme has situational \( S^d \) (i.e. presentation of materials is \( S^d \)) then count "a" as N/A.

Instructional Stimulus -

a. Simple - Written at appropriate language level for the child.
b. Consistent with goal behavior - if goal behavior is to get dressed upon the command "get dressed", that the instruction is "get dressed", not "put on your socks" for one step and "put on your shirt" for another step. These other verbalizations should be listed as prompts.

Prompt -

NOTE: If prompts not required then "a,b,c" are N/A.

a. Behavioral description of prompt type i.e. "hand on hand" rather than "manual prompt".
b. Prompts are faded in a consistent manner throughout the programme, appropriate to functioning level of the child.
c. Specification of the timing of the prompt, i.e. upon completion of steps 1 to 4 prompt for step 5.
Response -

a. Stated in behavioral, measurable terms.
b. The time limit for the response completion is stated.
c. If the time criteria is to change, that it is done consistently, i.e.: 
   1 minute for step 1 - putting on socks
   3 minutes for step 3 - putting socks, underwear and pants
   5 minutes for mastery - getting completely dressed

d. As the topography for intensity (not duration as that is evaluated in C) of the behavior changes toward the target behavior, that it is done consistently throughout the programme steps.

Consequences/Schedule -

a. Positive Reinforcement specified for correct behavior.
b. Punishment is specified for incorrect behavior.
c. Schedule of reinforcement for correct behavior given.
d. Schedule of punishment for incorrect behavior given.

Criterion -

a. Pass criteria for each step given.
b. Fail criteria for each step given.
c. Branching or recycling step indicated for fail criteria, or system described i.e. if fail criteria met for steps 1b - 1e recycle to last step. (Optional). If not given count as N/A.
d. Pass criteria for probes given.
e. Fail criteria for probes given.
Scoring Key Definitions - DTF
Page 5

Data Collection -
a. Indicate general method i.e. trial x trial time sample

b. Indicate specific notations used.

Others -
a. Probes are specifically notated in programme steps. (Optional if listed on face page).

+ = correct/complete
- = incorrect/not complete
N/A = not applicable for programme/or not taught

DL: jaor
16.7.80
DTF SCORING SHEET

INSTRUCTOR'S NAME ____________________________ TAPE # _____ START _____ END _____ PROGRAM INFORMATION ______

CHILD'S NAME ____________________________ DATE OF VIDEO ____________________________

TRIALS

I. PREPARATION:
   A. Fails to get attention
   B. Improper presentation of R.A.(s)

II. INSTRUCTIONAL STIMULUS:
   A. Fails to assure attention
   B. I.S. is not distinct
   C. I.S. is inconsistent/incomplete

III. PROMPT:
   A. Prompt is inconsistent with prog.
   B. Prompt is improperly timed

IV. RESPONSE:
   A. R. is inconsistent with prog. def.
   B. R. is inconsistent with prog. tim.

V. CONSEQUENCE:
   A. Not easily discernable
   B. Not immediate
   C. Inconsistent/ambiguous

VI. INTERTRIAL INTERVAL:
   A. Fails to remove materials
   B. Fails to record data
   C. Lack of distinct off-set

COMMENTS:

DL. 6.8.80

Total possible ______
- Total incorrect ______
= Total correct ______ Percent score ____%
EVALUATION FORM KEY

DISCRETE TRIAL FORMAM TEACHING

SECTION I: PREPARATION

A. Preparatory attention:

1. Child is sitting still
2. Child has hands on table
3. Child is looking at teacher or general work area.
4. Child is quiet

(N.B. Problems of judging whether or not the child is attending or not are enormous with some children: often the correctness of the child’s response or the assessment of the individual instructor of the readiness of the child are the only valid criterion the evaluator can use to decide if the child is attending.)

B. Proper presentation of response alternatives:

1. The instructor gives no external cues when presenting response alternatives.
2. The instructor does not obstruct the child’s view of the response alternatives.
3. The instructor does not use the response alternatives to gain the child’s attention; e.g., does not wave the R.A. (s) in the child’s visual field in order to gain or gage the child’s attention.

SECTION II: INSTRUCTIONAL STIMULUS

A. Child attends (see section I-A)

B. Instructional stimulus is distinct:

1. Isolated; e.g., there is a pause between everything preceding, specifically the presentation of the response alternatives or any attempts on the part of the instructor to regain the child’s attention.

2. No extra verbalizing, coaching on the part of the instructor, e.g., Do not say the child’s name after the I.S. is given.
C. Instructional stimulus is consistent and complete:

1. The instructional stimulus is always presented exactly as it is outlined in the program. (It is alright to randomize the instruction if the program is one of generalization).

Note: In the event that the $S^d$ is non verbal, such as the presentation of materials to be collated being the $S^d$ to work, the instructional stimulus section is counted as non applicable i.e. N/A. This would also be the case for a program written for free operant eye contact.

SECTION III: THE PROMPT (optional)

Rule: The prompt is defined as any teacher behavior related to the task or the student which occurs between the Instructional Stimulus being given and the student's first movement to respond. Thus correction procedures are counted in the consequence section.

EXCEPTION: If the program is a forward chaining procedure the prompt will occur after the beginning of the student's response. In this case the teacher's behavior would be counted as a prompt. For this to be the case however, you must have verification from the teacher that they are doing a forward chaining procedure.

A. The prompt is consistent:

1. Prompt is always presented and faded as outlined in program.

2. The instructor uses no extraneous cues and prompts.

B. The prompt is properly timed:

1. The prompt comes between the Instructional Stimulus and the response. For the prompt to be properly timed it must also come before the end of the time period allotted for the response. To distinguish between late prompts, those that come after the response and a correction, if the behavior was used as a correct prompt in the past or is written in the program as a prompt, then mark it as incorrect in the prompt box. If it is any other behavior it is marked in the consequence box.

SECTION IV: THE RESPONSE

A. The response is made and consequenced in complete accordance with the program definition. (Remember this is meant to evaluate the trainer's use of the behavior definitions for correct and incorrect responses not whether the child responded correctly or not).

B. The response is timed within the boundaries outlined in the program definition. e.g. If the response definition states that the response is to occur within 10 sec. of the instruction that the trainer uses this as a criteria for judging the correctness of the response.
Note: Thus if the teacher punished or extinguished a correct response then an error is marked in the response box. If they reward an incorrect response, again an error is marked in the response box. Of course, if either of these situations occurs, it will also mean the teacher will be marked wrong in the consequence section for giving an ambiguous consequence. As can be seen then, all errors in the instructor's use of the response definition will lead to an error also being marked in the consequence section. This is justified, since the situation leads to very bad teaching behavior.

SECTION V: THE CONSEQUENCE

A. The consequence is easily discernible:
   1. It must be heard or seen clearly on tape.

B. The consequence must be immediate:
   1. Unless specified the consequence should be given between 1-2 seconds upon completion of the response, or sooner.

C. The consequence must be consistent and unambiguous:
   1. It must follow exactly the definition outlined in the program.
   2. It must be clearly related to the response.
   3. The instructor should face the child when delivering a positive consequence, or verbal reprimand. If the trainer is ignoring the child, it must be obvious on tape.

RULE: All correction behavior of the instructor is counted in the consequence section. Correction behavior is defined as any instructor behavior towards the materials or the student which follows the beginning of the student's response, and is aimed at changing the child's immediate response. The exception to this definition is during a forward chaining procedure (see Section III Note and Exception). Unless designed as part of the teaching program (written in the program) all correction behavior of the instructor is marked as incorrect (item C - consequence is not consistent with program). If there is an acknowledged, structured, correction procedure, in the program and it is done correctly then it is evaluated as being correct in the consequence section.

SECTION VI: THE INTERTRIAL INTERVAL

A. The instructor removed materials immediately following the delivery of the consequence.

B. The instructor must record data after each trial.
C. The trial must have a distinct offset consisting of a 2-5 second pause before initiating a new trial. However, if the program calls for massed sequential trials as in collating stacks of papers, where the S is the materials being present, (see Section II-D) then no distinct offset will occur. This will be true for any program where independent work skills are being taught.

(N.B.: Reinforcement of on-task behavior during the intertrial interval is optional and will not be evaluated.)

D. The events of the intertrial interval should follow in this sequence:

1. Remove materials
2. Records data
3. Reinforcement of on-task behavior (optional)
4. Pauses to give trial a distinct offset.

DL:jaor
6.8.80
We would like to have some measure of our effectiveness. Please rate your responses to the following statements with numbers 1 to 5.

5 = Highest, excellent, strongly agree.
4 = High, good, mildly agree.
3 = Average, O.K.
2 = Low, not so good, mildly disagree.
1 = Lowest, really bad, strongly disagree.
N.S. = Not sure.
TRAINEE FEEDBACK SHEET

1. The preparation for the workshop was good.

2. The organization of the workshop was good.

3. The content of the workshop met my expectations.

4. The first day's introduction helped me understand what P.I.R.T. was, and what to expect.

5. I liked and learned from the following presentations:
   a. why D.T.F.?
   b. Observation Interpretation.
   c. Behavioral Principles.
   d. D.T.F. Components.
   e. Pre-requisite Skills.
   f. Baselining and Assessment.
   g. Programme writing lecture.
   h. D.T.F. vs Free Operant.
   i. Prompts, Probes, Branches, Maintenance, Generalization and Criteria.
   j. Communication Issues and Language Teaching.
   k. Techniques of Behavior Management.
   l. Task Analysis.
   m. Curriculum.
   n. Token Economy.

6. I liked and learned from the films.

7. I liked and learned from the following components of "hands on" training:
   a. Demonstrations (watching others).
   b. Role playing (practicing with adults).
   c. Hands on (practicing with a child).
   d. Video-tape feedback.
   e. Small group planning.

8. I thought the homework and feedback was relevant to my needs.

9. I had many opportunities to ask questions and talk about the workshop topics.

10. Please comment on at least three of the above statements. For example, why you liked or didn't like a presentation. (Please use back of page if space is not sufficient.)
APPENDIX G
Inservice Resource Teams

Name: ____________________________

FOLLOW-UP FEEDBACK #1

As this is our final follow-up, we would like to gather some data regarding our effectiveness at helping you apply the information and techniques discussed during the workshop and follow-ups to your day-to-day work situation.

Please note that any questions asking for a quantity such as how often you have done a specific task should be filled in with a specific number. For several of the questions, it may be helpful to consult your records.

Please make sure that you have signed your name to all the sheets in this form.

In advance, we would like to thank you for giving the time and energy required to complete this form. Your effort is appreciated.

W.I.R.T./P.I.R.T.
Inservice Resource Team

FOLLOW-UP FEEDBACK #1

1. Are you using DTF Programme Plans similar to those taught at the workshop?  Yes  No

1a. If yes, how many programmes are you using? 

1b. If no, why not?

2. Are you using data collection techniques learned during the workshop?  Yes  No

2a. If yes, how many behaviors are you keeping data on?

2b. If no, why not?

3. Are you using a Discrete Trial Format for any of your teaching?  Yes  No

3a. If yes, for how many tasks are you using it for?

3b. If no, why not?

4. Have you written a DTF Programme Plan such as those done during the workshop within the last 6 months?  Yes  No

4a. If yes, how many have you written?

4b. If no, why not?
5. Do you baseline a child before teaching a new behavior?
   Yes ____  No ____

5a. Do you baseline a child before you start a behavior management programme? Yes ____  No ____

6. Are you using an assessment technique that requires you to take objective data of the behaviors being assessed? Yes ____  No ____

7. If you were to do the tasks listed below, how confident would you be that you could now do them independently?

<table>
<thead>
<tr>
<th>not at all</th>
<th>hesitant</th>
<th>O.K.</th>
<th>Confident</th>
<th>Very Confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Teaching using DTF at an 80% correct criterion.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>b. Writing a DTF Programme Plan at an 80% correct criterion.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>c. Keep data and graph it.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>d. Use a token economy.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>e. Build an Individual Student Curriculum.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>f. Use the Behavior Management Techniques taught in the workshop.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>g. Teach someone else to use DTF</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>h. Teach someone else to use behavior management techniques.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
8. Think back to before the workshop. Now do question 7 again. This time rate how confident you were before the workshop at doing these tasks independently.

<table>
<thead>
<tr>
<th></th>
<th>not at all confident</th>
<th>hesitant</th>
<th>O.K.</th>
<th>confident</th>
<th>very confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Teaching using DTF at an 80% correct criterion.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>b. Writing a DTF Programme Plan at an 80% correct criterion.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>c. Keep data and graph it.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>d. Use a token economy.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>e. Build an Individual Student Curriculum.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>f. Use the Behavior Management Techniques taught in the workshop.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>g. Teach someone else to use DTF.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>h. Teach someone else to use behavior management techniques.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Inservice Resource Teams

Name: ____________________________

FOLLOW-UP FEEDBACK #2

As this is our final follow-up, we would like to gather some data regarding our effectiveness at helping you apply the information and techniques discussed during the workshop and follow-ups to your day-to-day work situation.

For those questions asking for a rating of 1-5, please circle the appropriate number. Please make sure that you have signed all the sheets in this form.

All comments will be appreciated. Thank you for your co-operation.

W.I.R.T./P.I.R.T.
Follow-up Feedback #2
Page 1

1. How useful do you feel the content of the workshop has been for you over the last six months.

No help at all                                      Very helpful

1  2  3  4  5

2. How much more did the information you received during the follow-up visits add to that of the workshop?

No help at all                                      Very helpful

1  2  3  4  5

3. How effective were we in helping you integrate our input into practical use in your teaching situations?

Not effective at all                                  Very effective

1  2  3  4  5

4. How much would you attribute any progress of the children to input which we have given to you?

Hindered progress                                  Facilitated progress

Progress the same

1  2  3  4  5

5. Were there any children you think our input was helpful for?

Yes ____ No ____ Please explain. ____________________________

________________________________________________________________________

________________________________________________________________________

6. Were there any children you felt our input was a hindrance to?

Yes ____ No ____ Please explain. ____________________________

________________________________________________________________________

________________________________________________________________________
7. How useful to you would it be to have P.I.R.T./W.I.R.T. come back on a consulting basis every 6-8 months.

<table>
<thead>
<tr>
<th>No use at all</th>
<th>Very useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

8. Do you have any further general comments about our follow-up visits?

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________

9. Do you have any further comments about the team in general?

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
Reliability

Percentage of scores checked and percentage agreement per test period.

### Terminology Measure

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Mid</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of scores checked</td>
<td>35</td>
<td>35</td>
<td>48</td>
</tr>
<tr>
<td>Mean percent agreement</td>
<td>93.9</td>
<td>99.1</td>
<td>97.2</td>
</tr>
<tr>
<td>Range</td>
<td>89-100</td>
<td>94-100</td>
<td>94-100</td>
</tr>
</tbody>
</table>

### DTF Measure

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Mid</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of scores checked</td>
<td>33</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td>Mean percent agreement</td>
<td>90.2</td>
<td>92.1</td>
<td>97.00</td>
</tr>
<tr>
<td>Range</td>
<td>80-100</td>
<td>92-100</td>
<td>83-100</td>
</tr>
</tbody>
</table>

### Program Writing Measure

<table>
<thead>
<tr>
<th></th>
<th>Mid</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of scores checked</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>Mean percent agreement</td>
<td>86</td>
<td>89</td>
</tr>
<tr>
<td>Range</td>
<td>79-95</td>
<td>77-95</td>
</tr>
</tbody>
</table>