4-1980

A Multiple-Baseline Analysis of Social/Academic-Skills Training Across Different Stimulus Conditions: Generalization as a Function of Stimulus Diversity

Stanley R. Carlock
Western Michigan University

A MULTIPLE-BASELINE ANALYSIS OF SOCIAL/ACADEMIC-SKILLS TRAINING ACROSS DIFFERENT STIMULUS CONDITIONS: GENERALIZATION AS A FUNCTION OF STIMULUS DIVERSITY

by

Stanley R. Carlock

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
April 1980

ACKNOWLEDGEMENTS

A debt of thanks is owed to a host of people, all of whom have contributed in one way or another to my successful completion of this thesis. The entire staff of Valley Center, namely, Monte Bender, Tim McKinley, Tom Noblet, Mike VanVaerenbergh, Mike Zaborowski, and Mark Zulauf, are all owed a great deal of thanks for their suggestions and cooperation in helping me obtain various types of data and releases. A special appreciation is due to teachers Mary Ashby and Beth Shinar for their patience, understanding, and tireless cooperation, and to Cyndee Brule for her excellent skills as a typist in producing the final copy of this paper. Special consideration and thanks are due to Jim Kaye, who has provided guidance, support, feedback, and access to equipment and facilities, without which this study would not have been possible. I would also like to personally thank Rich Jackle for his time and suggestions in helping me obtain reliability data, as well as Dawn Worley, who not only helped me obtain reliability data, but who also provided a great deal of support and understanding throughout the course of my studies at Western. My final thanks are due to the members of my program committee: Drs. Jack Michael, Wayne Fuqua, and Richard Malott. All of whom have imparted to me a greater understanding of theoretical, experimental, and applied behavior analysis through their courses, and private and public consultation. In particular, I would like to express a special appreciation to Professor Michael for allowing me a great deal of freedom in pursuing the area of research in which I am particularly interested.

Stanley R. Carlock

1314831

CARLOCK, STANLEY RAY
A MULTIPLE-BASELINE ANALYSIS OF SOCIAL/ACADEMIC-SKILLS TRAINING ACROSS DIFFERENT STIMULUS CONDITIONS: GENERALIZATION AS A FUNCTION OF STIMULUS DIVERSITY.

WESTERN MICHIGAN UNIVERSITY, M.A., 1980

University Microfilms International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

INTRODUCTION

In 1977, Stokes and Baer described nine general "types" of generalization, one of which involved the training of sufficient exemplars. The notion behind this aspect of generalization training was... "if the results of teaching one exemplar of a generalizable lesson is merely the mastery of the exemplar taught, with no generalization beyond it, then the obvious route to generalization is to teach another exemplar of the same generalization lesson, and then another, and then another, and so on until the induction is formed (i.e., until generalization occurs sufficiently to satisfy the problems posed)." Indeed, the primary concept upon which the training of sufficient exemplars is based, is that generalization to untrained stimulus conditions, and to untrained responses is achieved by the training of "sufficient" exemplars (rather than all) of the desired stimulus conditions or responses.

While perhaps being one of the most valuable areas or programming, and also "the generalization-programming area most prominent and extensive in the present literature... very little research concerned with generalization programming has dealt with the training of sufficient stimulus exemplars" (Stokes and Baer, 1977). However, the training of sufficient exemplars mode of generalization training has been successful in training retarded children in verbal responses to a wider range of people in their environment (Garcia, 1974, and Stokes, Baer, and Jackson, 1974). Instances in which this type of generalization training has been successfully used to improve responding in a number of different settings have also been reported with deviant children (Allen, 1973), autistic children (Rincover and Koezgel, 1975), a retarded woman (Griffiths and Craighead, 1972), and retarded children (Murdock, Garcia, and Hardman, 1977, and Frisch and Schumaker, 1974). Increasing response class repertoires in retarded children (Clark and Sherman, 1975; Garcia, Baer, and Firestone, 1971; and Baer, Peterson, and Sherman, 1967), teachers (Horton, 1975), female psychiatric patients (Tracey, Briddel, and Wilson, 1974), and a male bilingual child (Bucher and Mueller, 1977) has also been successful using sufficient exemplar generalization training.

Two additional studies (Frederiksen, et al., 1976, and Bornstein, Bellack, and Hersen, 1977) which were indirectly concerned with the training of sufficient stimulus exemplars, focused (in particular) upon training social-skills in psychiatric patients, and unassertive children, respectively. Both of these studies presented their subjects with a minimum of six training situations (role-played scenes), with additional untrained scenes being presented in order to assess whether or not generalization from trained to untrained situations had occurred. Both studies were comprised of a multiple-baseline design, one across behaviors (Bornstein, Bellack, and Hersen, 1977), and the other across subjects (Frederiksen, et al., 1976). In neither study, however, were the results considered in terms of the "sufficiency" of their exemplars (role-played scenes) to the extent to which generalization occurred from training situations to untrained or in vivo situations.

The aim of the present study was to utilize a multiple-baseline analysis across stimulus situations in order to examine to what extent training on the dependent measures of particular stimulus situations would be sufficient to facilitate improvements in dependent measures of untrained situations (particularly, classroom situations). Two experiments were conducted: the first consisted of training social-skills in a subject, with a procedure which roughly approximated that described by Bornstein, Bellack and Hersen (1977), and Frederiksen, et al. (1976). The second experiment focused on training a subject in social/academic skills, also following the procedure discussed above. In this manner, it was also hypothesized that generalization, as a result of training in a number of rather diverse stimulus situations, could be achieved.

Experiment I

This experiment was initiated in an attempt to develop a more socially adaptive behavioral repertoire in a 14 year old, emotionally impaired male student. This student's teacher described him as exhibiting certain behavioral excesses; behaving in an overly "dramatic" and "serious" manner in response to a number of particular circumstances.

During training in social-skills, specific component behaviors were targeted toward which training was directed. In order to facilitate a more diverse repertoire, three rather different stimulus situations were presented, with training being focused upon improving (i.e., either increasing or decreasing) each behavioral component of the subject's overall set of responses.

The following experiment was designed to assess whether or not training in one stimulus situation resulted in generalization effects being evident in other stimulus situations, and particularly in the classroom setting.

METHOD

Subject

Subject-1 was a student in a self-contained classroom for emotionally impaired children and adolescents. He was referred by his teacher, who reported that when a number of particular situations occurred relevant to this student, his typical responses consisted of behaviors that were excessive in nature. In particular, when this student was given something (e.g., a snack, a compliment, extra points for staying on task, etc.), saw someone (a favorite substitute teacher or aide) whom he interacted with only on an intermittent basis, or was confronted with certain issues (not having completed his homework, not knowing an assignment, teasing other people, etc.), he responded in a manner which was excessively overdone (i.e., too "dramatic" or "serious") (e.g., a typical response to being given a snack might be: "Oh thank you, how nice you are for doing this. What a thoughtful thing to do."). The student also frequently attempted to avoid or escape certain situations by changing the subject (e.g., When being confronted with not having completed an assignment, the subject's response might be: "Can I sharpen my pencil now?"). This type of responding occurred within the classroom or school setting on a frequent basis, and typically continued anywhere from one response to several interchanges (between the student and the other person(s) involved). This subject's interactions with other people usually tended to be appropriate, although he was somewhat overly polite, and teased his peers on occasion. This subject was also enrolled in two hours of classes (one regular education class and one special education class) at a local junior high school.

In gaining informed consent to participate in this study, subject-1 and his parents were informed of the research nature of the program, as well as their volunteer status. It was explained that that study's goal was to increase the interpersonal skills of the subject. Assessment began upon obtaining informed consent from the subject and his parents.

Setting and Equipment

All sessions were conducted in a small corner of an unoccupied classroom which had been converted to accomodate highly disruptive students under exclusion time-out procedures. Three sessions were conducted per day, five days a week, for thirty minutes. At the beginning of each session, the subject was seated at a card table across from the investigator. Immediately behind and to the right of the investigator, a portable Panasonic video-camera was placed in order to obtain a concise video-recording of the subject's responses. A remote microphone was suspended from the ceiling just above the table so that clear sound reproduction could be achieved. Other equipment included a portable Sony video-recorder and monitor used for playback during videotape data extraction, and a standard stopwatch for measuring percentage of eye-contact.

PROCEDURE

Selection and Scoring of Target Behaviors

A behavioral analysis of the subject's responses was conducted via classroom observations and interviews with his teacher. Based on this analysis, four dependent measures were selected and are defined below.

Ratio of eye-contact to speech duration. The total length of time in seconds that the subject looked at the investigator (or person being interacted with) while he (the subject) was speaking was measured for each situation. The ratio was computed by dividing the total duration of eye-contact while speaking by the total duration of speech.

Appropriate comments. Verbal responses which were acceptable, relevant to the given situation, and which did not indicate direct or indirect psychological or physical harm (e.g., "Let's check with someone about the game rules.").

Irrelevant comments. Verbal responses which were unrelated to the theme of a given situation, and indicated an attempt at avoidance or escape from the situation (e.g., "I'm sorry, we'll play the game your way.").

Level of affect. The subject's affect was scored on a five-point scale (1 = a very inappropriate or inexpressive tone of voice and absence of appropriate facial and physical gestures; 5 = a full and appropriate tone of the voice, with corresponding facial and physical gestures appropriate to each situation).

Behavioral Assessment

Role-played scenes. Three role-played vignettes were developed for this subject which corresponded and were highly similar to situations which prompted behaviorally-excessive responses. An attempt was made to include situations that the subject was likely to engage in daily with other persons. Listed below are the three vignettes of which the role-play situations were comprised.

Situation-1.

Scene: It's Monday morning. You walk into the classroom after getting to school. There you see Mrs. _____ at the back of the room making hot chocolate. She looks up to see you standing there, and begins to say good-morning.

Prompt: "Good-morning _____, it's Monday morning and I wanted to start the week off by doing something special. Here, have a doughnut and a cup of hot chocolate."

Situation-2.

Scene: During rec. you start up a game of pool with _____. You make a shot that knocks one of your balls in, then aim and take another shot. _____ then accuses you of knocking one of his balls in, and of taking an extra shot.

Prompt: "Hey _____, you knocked one of my balls in, why'd you take another shot?"

Situation-3.

Scene: Mrs. _____, one of your favorite substitute teachers, whom you haven't seen in a few weeks, meets you at the door when you come in.

Prompt: "Good morning _____, nice to see you. How are you doing this morning?"

The initial (baseline) assessment consisted of seven presentations of each stimulus situation, following instructions to the subject to respond as if he were actually faced by the situation described. Following each scene description, the prompt was delivered. After the subject responded to the prompt, a second statement (prompt) was made by the investigator (relevant to the situation presented) which was followed by a second response from the subject (two complete interchanges). Subsequent scenes were administered in the same fashion without the subject receiving any feedback, instructions, or modeling from the investigator.

Classroom probes. Observations were made of situations which occurred in the classroom involving Subject-1 and another person. Situations were chosen which best approximated instances that elicited responses of interest, with the observer scoring all dependent measures in accordance with the definitions described above. During baseline assessment, one classroom probe observation was carried out, with three occurring during training, and two during the follow-up phase.

Training

During each session, Subject-1 received training on each of the three vignette scenes. The training procedure consisted of the following components: (a) The investigator presented the first scene, delivered a prompt, obtained a response from the subject, replied to the subject's response, and obtained a second response from the subject. (b) The investigator then provided the subject with feedback on his performance, with reference to each specific target behavior (i.e., the investigator explained to the subject the relative appropriateness of his responses to the situation presented). (c) The investigator then discussed feedback with the subject to entire that he understood (i.e., The subject was asked to explain his understanding of how he had responded, and how he might respond more appropriately to future situations). (d) The subject and the investigator then reversed roles, with the investigator modeling responses, giving special attention to target behaviors (i.e., The subject presented the prompt, followed by the investigator modeling appropriate responses). (e) Specific instructions were then given by the investigator concerning the target behaviors (i.e., The subject was requested to respond in ways approximating the responses modeled by the investigator), followed by having the subject respond to subsequent situations, and then again to the trained scene. (f) Rehearsal then continued for a scene until the investigator believed that the criterion for the target behaviors had been achieved. (g) Training then advanced to the next untrained situation, proceeding in a similar fashion through all training scenes.

Consistent with multiple-baseline strategies, training was directed sequentially and cumulatively across the three vignette scenes over the 24 session period. Following baseline, the subject received training on situation-1 over six sessions. Training on situation-2 extended over the next nine sessions; situation-3 also required nine training sessions. In addition, follow-up probes were obtained at 1 and 2 week intervals after training. Concomitantly, classroom probes were conducted at intervals extending throughout baseline, training, and follow-up.

Reliability of Observations

Percentages of agreement for each dependent measure are summarized in Table 1.

Role-played scenes. Two judges independently rated all videotapes for Subject-1 retrospectively. For the measure of percentage of eye-contact, agreements were scored if the independently obtained percentages did not differ by more than $\pm$ 4%. Discrepancies of more than 4% were scored as disagreements. Likewise, for the measure of level of affect, agreements were scored if the independently assigned ratings did not differ by more than 10%. Discrepancies of more than 10% were scored as disagreements.

For all dependent measures, agreement was calculated using two methods to ensure that reliability was acceptable for both high-frequency and low-frequency behaviors. In the first method, agreement was scored only if both judges agreed that the behavior had either occurred or had not occurred during a particular scene. Disagreement was scored if only one judge scored the occurrence of a behavior. Per cent agreement was calculated by dividing the total number of agreements plus disagreements x 100. Agreement was recalculated without counting the number of agreements on nonoccurrence. Therefore, the second set of agreement percentages in Table 1 is based on the number of agreements on occurrence over the total number of agreements plus disagreements.

Table 1

Percentage agreement calculated by two methods on all dependent measures for Subject-1.

Table 1

Dependent Measure	Method: Including Agreement on Nonoccurrence	Method: Excluding Agreement on Nonoccurrence
Role-Played Scenes:		
Percentage of eye-contact	84.5	
Percentage of appropriate comments	93.3	81.2
Percentage of irrelevant comments	96.3	89.1
Level of affect	88.6	
Classroom Probes:		
Percentage of eye-contact	81.1	
Percentage of appropriate comments	98.4	92.7
Percentage of irrelevant comments	100.0	100.0
Level of affect	92.3	

Classroom probes. One judge rated all presentations of classroom situation probes. Two additional judges independently rated one-third of these presentations. The same methods for calculating agreement for all dependent measures as were used on the role-played scenes were performed. These percentages of agreement for each dependent measure are also summarized in Table 1.

RESULTS AND DISCUSSION

Figures 1, 2, and 3 show the results of the multiple-baseline analysis of the four dependent measures for Subject-1 across baseline, training, and follow-up. While baseline was in effect for situation-1 (Fig. 1), a considerable amount of variability was observed for all dependent measures across all situations, with a certain degree of stability occurring in some instances (particularly for percentage of eye-contact for situation-1, percentage of appropriate comments, and percentage of irrelevant comments for situation-2, and level of affect for situation-3). When the training phase was implemented for situation-1, a considerably rapid and maintained improvement was noted for practically all measures (in situation-1), level of affect being the exception, where little change was noted (maintaining between a level of 3 and 4). Generalization effects across any dependent measures for situations-2 and 3 were not noted during training on situation-1. When the training phase for situation-2 (Fig. 2.) was implemented, a fairly rapid improvement was again noted for percentage of appropriate comments (increased from 50% to 100%), and percentage of irrelevant comments (decreased from 50% to 0%) (in situation-2). Percentage of eye-contact appeared to stabilize at a fairly high level (approximately 100%), while level of affect appeared to improve somewhat (from an average level of 3 to an average level of about 4.25), and gained a small amount of stability. In terms of generalization effects, percentage of appropriate comments, and percentage of irrelevant comments (in situation-3) (which appeared to begin improving dur-

Figure 1.

Sessions across baseline, training, and follow-up for Subject-1. A multiple-baseline analysis of: percentage of eye-contact, percentage of appropriate comments, percentage of irrelevant comments, and level of affect across the first stimulus situation.

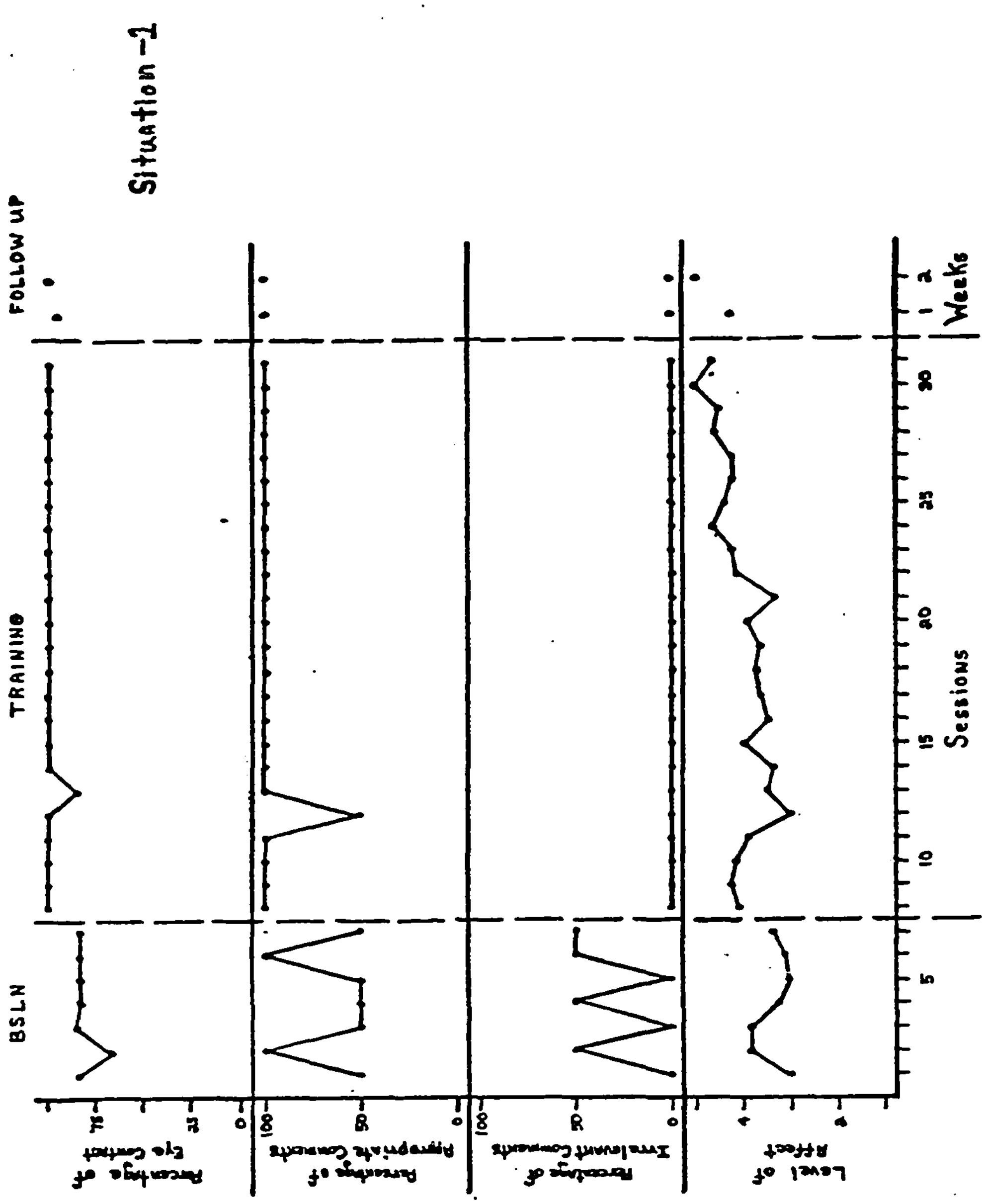
Situation -1
BSLN
TRAINING
FOLLOW UP
Percentage of Eye Contact
Percentage of Appropriate Comments
Percentage of Irrelevant Comments
Level of Affect
Sessions
Weeks

Fig. 1.

Figure 2.

Sessions across baseline, training, and follow-up for Subject-1. A multiple-baseline analysis of: percentage of eye-contact, percentage of appropriate comments, percentage of irrelevant comments, and level of affect across the second stimulus situation.

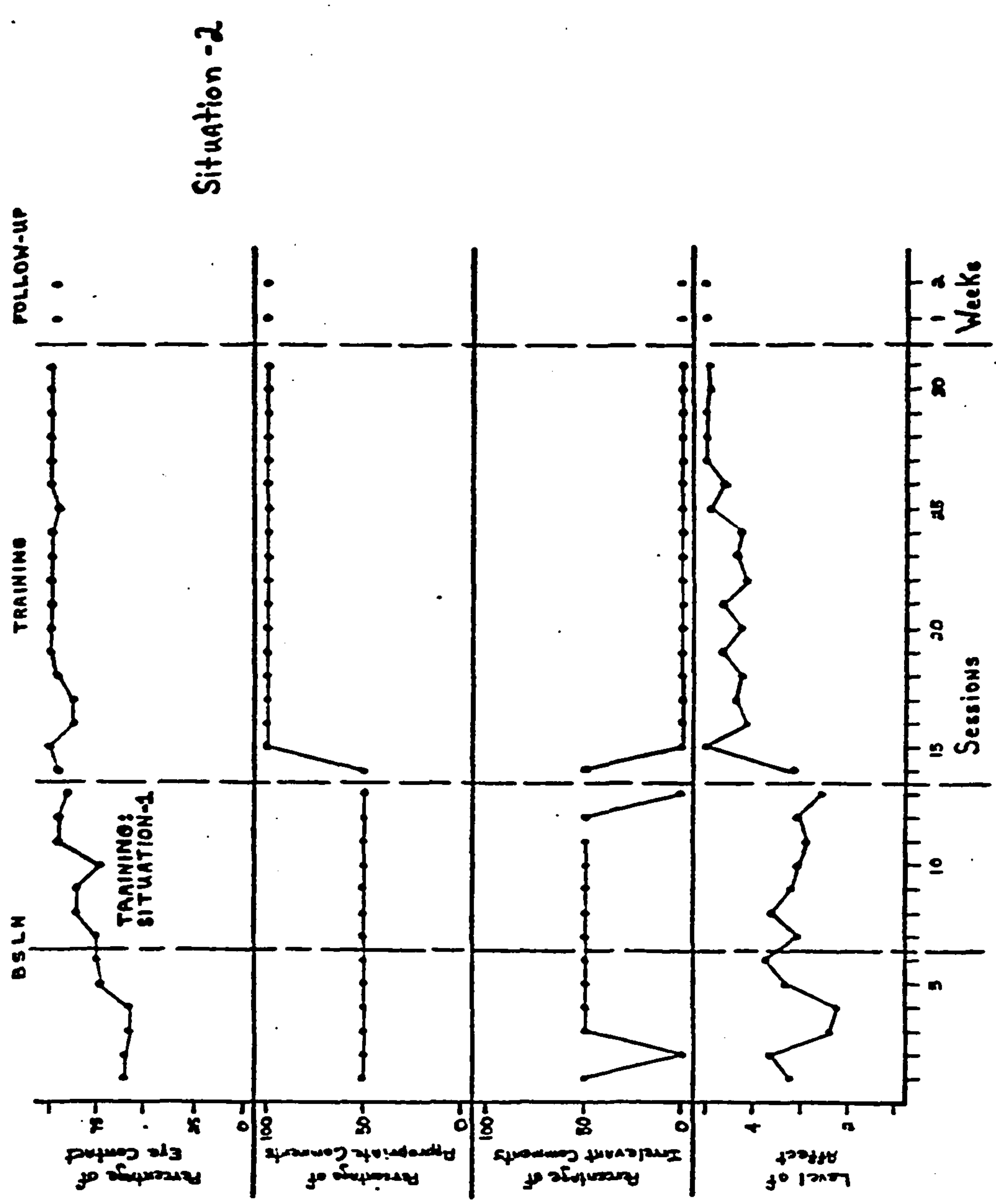
Situation -2
BSLN
TRAINING
FOLLOW-UP
TRAINING: SITUATION-1
Percentage of Eye Contact
Percentage of Appropriate Comments
Percentage of Irrelevant Comments
Level of Affect
Sessions
Weeks

Fig. 2.

Figure 3.

Sessions across baseline, training, and follow-up for Subject-1. A multiple-baseline analysis of: percentage of eye-contact, percentage of appropriate comments, percentage of irrelevant comments, and level of affect across the third stimulus situation.

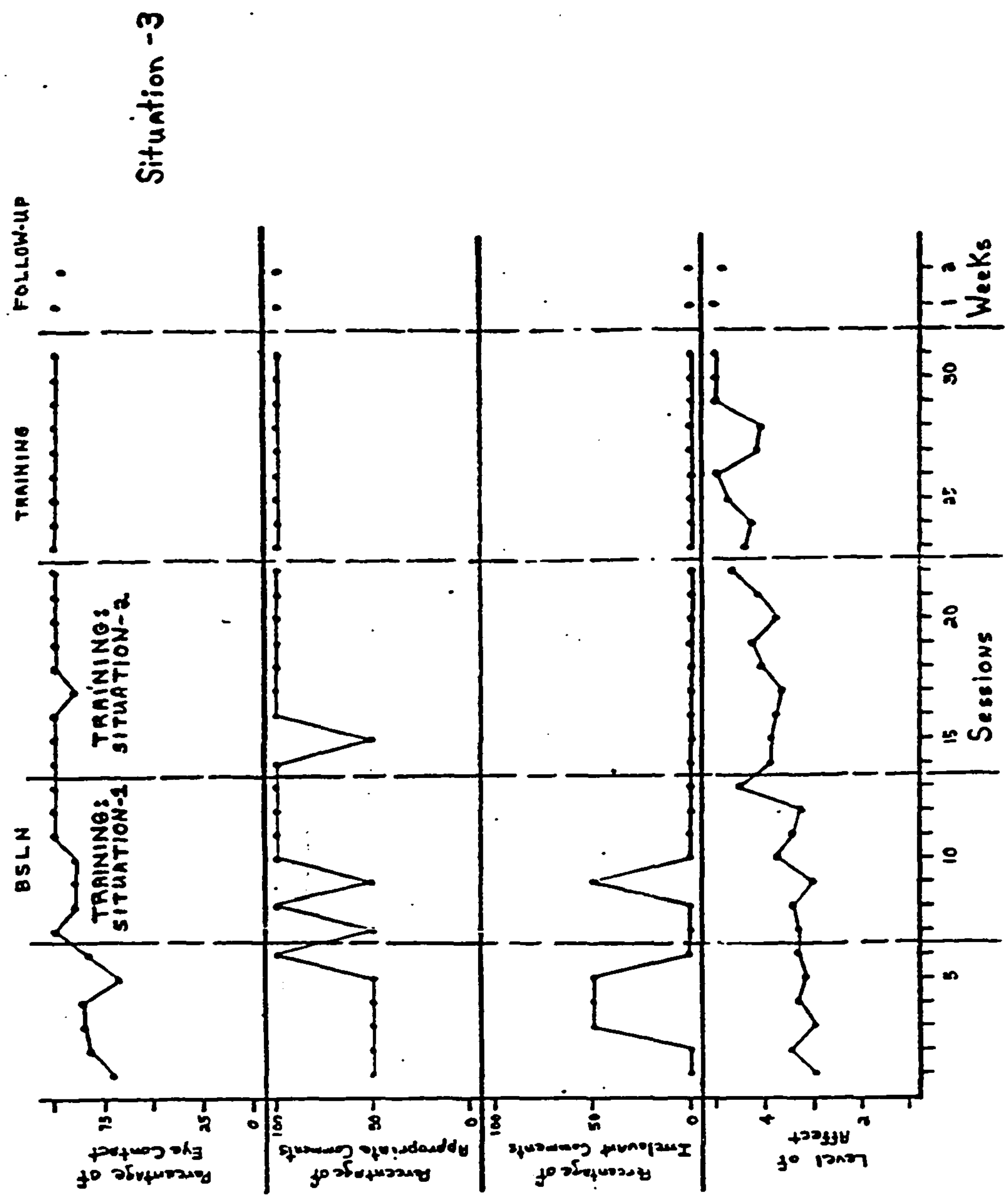
Situation -3
BSLN
TRAINING: SITUATION-1
TRAINING: SITUATION-2
TRAINING
FOLLOW-UP
Percentage of Eye Contact
Percentage of Appropriate Comments
Percentage of Irrelevant Comments
Level of Affect
Sessions
Weeks

Fig. 3.

ing training on situation-1) stabilized after achieving a desirable rate in their respective directions (100% and 0% respectively).

When training on situation-3 (Fig. 3.) was implemented, some change was observed for all of the dependent measures (in situation-3), with the exception of a slight increase in level of affect (up from an average level of about 4 to an average level of about 4.5, although this measure had been gradually improving since training had begun, during situation-1). Generalization effects were also minimal, with some apparent improvement and stabilization however, in level of affect for both situations-1 and 2.

Follow-up observations at one and two weeks indicated that little change in the subject's overall set of responses had occurred since the end of the training phase. Across all three situations, the dependent measures appeared to be maintained at a desirable level.

Table 2 shows the average percentage rate of responses for Subject-1, taken from classroom probe observations (data were collapsed due to the small number of observations: 1 during baseline, 3 during training, and 2 during follow-up). Some degree of improvement was noted for each dependent measure, with the exception of percentage of irrelevant comments, which was at zero during baseline (although it returned to zero during follow-up). In most cases, follow-up rates were better than rates at the end of training.

The results of this multiple-baseline analysis indicate that implementation of the training phase coincided with a noticable

Table 2

Average percentage rate of responses from classroom probe observations for Subject-1.

Table 2

Dependent Measure	Experimental Phase		
	Baseline	Training	Follow-Up
Percentage of eye-contact	55.0	63.0	91.0
Percentage of appropriate comments	50.0	100.0	100.0
Percentage of irrelevant comments	0.0	33.3	0.0
Level of affect (not a per cent)	3.8	4.0	4.3

improvement in most of the dependent measures, which was not only rapid in onset, but maintained well in all of the role-played situations, and continued into follow-up observations at one and two weeks. Although widespread generalization effects were not noted, a certain amount of generalization seemed evident, particularly for percentage of appropriate comments, and percentage of irrelevant comments in situation-3, while training was in effect for situations-1 and 2. And although classroom probe data were scant, a general improvement seemed to have taken place, particularly, after the training phase had ended (i.e., during follow-up). Anecdotal reports from the subject's teacher also seemed to substantiate a general improvement in the subject's responses to situations in the classroom (which had previously elicited maladaptive responses).

Experiment II

This experiment was initiated in an attempt to promote the development of a combination of social and academic behaviors in a 16 year old, emotionally impaired, male student. This student's teacher described him as exhibiting certain behaviors which were imcompatible with doing well in the classroom, in terms of interactions between his teacher and himself, and achieving acceptable grades on his academic work.

During training, specific behaviors were targeted toward which training was directed. Three different sets of stimulus conditions were developed in order to approximate a wide array of social/academic conditions which typically occurred in the classroom setting. Training then focused upon improving (i.e., increasing or decreasing) each target behavior comprising the student's overall set of responses.

The following experiment was designed to assess whether or not training in one stimulus situation resulted in generalization effects being evident in other stimulus situations, and particularly, in the classroom setting.

METHOD

Subject

Subject-2 was a student in a self-contained classroom for emotionally impaired children and adolescents. He was referred by his teacher, who described him as behaving in a "hyperactive" manner. In particular, he was frequently known to interrupt his teacher when instructions were being presented relevant to academic assignments. In addition, he frequently would begin working on an assignment before his teacher had completed her explanation of the instructions, and could rarely correctly repeat the instructions which he had been given. Consequently, he received comparatively low grades on his academic work.

Although this subject resided in the children's unit of a local psychiatric hospital, his interactions with staff and peers in the classroom were considered as being relatively appropriate. An exception existed however, with the student's younger brother, who was enrolled as a student in the same classroom as Subject-1 just after the onset of this experiment (sessions 12-14,see Figures 7, 8, and 9). At that point, the subject was observed to exhibit a somewhat higher rate of disruptive (i.e., "hyperactive") behavior than before his brother was enrolled as a student.

The subject had a history of receiving psychotropic medications, and underwent two alterations in prescribed drug therapies during the course of this investigation (he was placed on Thorazine, 25 mg., b.i.d. between sessions 14 and 15, and then changed to

Ritalin, 10 mg., b.i.d. during between sessions 29 and 30).

In gaining informed consent to participate, Subject-1 and his parents were informed of the research nature of the program, as well as their volunteer status. It was explained that the goal was to increase the social/academic skills of the student. Assessment began upon gaining informed consent from Subject-2 and his parents.

Setting and Equipment

All sessions were conducted in a corner area of a small, unoccupied classroom which had been converted to accomodate highly disruptive students under exclusion time-out procedures. Three sessions were conducted per day, five days per week, for 30 minutes. At the beginning of each session, the subject was seated at a card table across from the investigator. Immediately behind and to the right of the investigator, a portable Panasonic video-camera was placed in order to obtain a concise video-recording of the subject's responses. A remote microphone was suspended from the ceiling just above the table so that a clear sound reproduction could be achieved. Other equipment included a portable Sony video-recorder and monitor used for playback during video-tape data extraction.

PROCEDURE

Selection and Scoring of Target Behaviors

A behavioral analysis of the subject's responses was conducted by way of classroom observations and reports by his teacher. Based on this analysis, four dependent measures were selected and are defined below.

Frequency of interruptions. An interruption was scored whenever the subject emitted a verbal response while the investigator was reading him the instructions (i.e., the subject "cut him off"), or after the subject had begun work on the problem set (instances in which the subject "mumbled to himself" while completing the problem sets were excluded).

Percentage of instructions repeated. The ratio of instructions which the subject correctly repeated (read) back to the investigator to the total number of instructions in a given problem set.

Percentage of instructions completed before beginning. The ratio of instructions read to the subject by the investigator to the total number of instructions which existed for a given problem set, prior to the subject beginning work on that problem.

Percentage of problems correct. The ratio of problems which the subject correctly completed to the total number of problems in a given set.

Behavioral Assessment

Stimulus problem sets. Three sets of stimulus conditions were developed for Subject-2 which corresponded and were highly similar to academic problems presented to the student in the classroom, and which typically lead to difficulties (e.g., interrupting his teacher, failing to allow instructions to be completed, being unable to repeat instructions correctly, and failure to correctly complete assignments) in the classroom. An attempt was made to include problems like those typically received as assignments on a daily basis in the classroom.

Figures 4, 5, and 6 represent the three sets of problems presented to Subject-2 (note that problems within each set varied from session to session). The initial baseline assessment consisted of six presentations of each set of stimulus conditions, following instructions to the subject to respond as he would to similar assignments in the classroom. Following the presentation of each problem, instructions were read to the student. Immediately afterward, the subject was requested to repeat the instructions and then was asked to complete the set of problems. Subsequent problem sets were then presented to Subject-2 in the same fashion, without the subject receiving any feedback, further instructions, or modeling from the investigator.

Classroom probes. Observations were made of assignments presented to Subject-2 in the classroom setting by his teacher or an assistant, with the observer scoring all dependent measures in

Figure 4.

Problem set-1, as presented to Subject-2 during baseline, training, and follow-up phases of this study.

Session ____________________ Date ________________

Problem Set-1

INSTRUCTIONS:

1) On the first line below, write your ____________ name.

2) On the second line below, write the name of the city ____________ live(s) in.

3) On the third line below, write the name of one of your ____________ friends.

4) Raise your hand when you are done.

(1) ________________________________

(2) ________________________________

(3) ________________________________

Fig. 4.

Figure 5.

Problem set-2, as presented to Subject-2 during baseline, training, and follow-up phases of this study.

Session ____________________ Date __________________

Problem Set-2

INSTRUCTIONS:

In the sentence below --

1) Draw a circle around all words which have the letter _____ in them.

2) Draw a line under all the words which have the letter _____ in them.

3) Draw a line through all the words which have the letter _____ in them.

4) Raise your hand when you are done.

Fig. 5.

Figure 6.

Problem set-3, as presented to Subject-2 during baseline, training, and follow-up phases of this study.

Session ______________________ Date ______________________

Problem Set-3

INSTRUCTIONS:

1) Divide _____ by _____, or $\overline{)}$

   Write your answer on the first line below.

2) Multiply _____ by _____, or x_____

   Write your answer on the second line below.

3) Subtract _____ from _____, or -_____

   Write your answer on the third line below.

4) Turn your paper over when you are done.

(1) ______________________

(2) ______________________

(3) ______________________

Fig. 6.

accordance with the definitions described above. During baseline assessment, one classroom probe observation was carried out, with three occuring during training, and two during the follow-up phase.

Training

During each session, the subject received training on each of the three problem sets. The training procedure consisted of the following components: (a) The investigator presented the first problem set to the subject, read the instructions for that problem set, and then asked the subject to read back the instructions. The subject was then allowed to complete the problem set (taking anywhere from approximately 30-90 seconds, depending upon the given problem set). (b) The investigator then provided the subject with feedback regarding his performance, with reference to each specific target behavior (i.e., The investigator explained the relative appropriateness of the subject's set of responses to the presentation of a given problem set). (c) The investigator then proceeded to discuss the feedback with the subject to ensure that he understood (i.e., The subject was asked to explain his understanding of how he had responded, and how he might respond differently in the future). (d) The subject and the investigator then reversed roles, with the investigator modeling responses, giving specific attention to target behaviors (i.e., The subject presented the prompt and the investigator modeled appropriate responses, without completing the problem set). (e) Specific instructions were then given by the investigator concerning the target behaviors (i.e., The subject was requested to respond in ways approximating the responses modeled by the investigator), followed by the presentation of subsequent problem sets, and then re-presentation of the trained problem set. (f) Rehearsal then continued for a given problem set until the investigator believed that the criterion for the target behaviors had been achieved. (g) Training then advanced to a new problem set, proceeding in a similar fashion through all sets.

Consistent with multiple-baseline strategies, training was directed sequentially and cumulatively across the three problem sets over the 26 session period. Following baseline, the subject received training on problem set-1 over five sessions. Training on problem set-2 extended over the next twelve sessions, with problem set-3 requiring nine training sessions. In addition, follow-up probes were conducted at one and two week intervals after training. Concomittantly, classroom probes were conducted at intervals throughout baseline, training, and follow-up.

Reliability of Observations

Percentages of agreement for each dependent measure are summarized in Table 3.

Stimulus problem sets. Two judges independently rated all video-tapes for Subject-2 retrospectively. For all dependent measures, agreement was calculated using two methods to ensure

Table 3

Percentage agreement calculated by two methods on all dependent measures for Subject-2.

Table 3

Dependent Measure	Method: Including Agreement on Nonoccurrence	Method: Excluding Agreement on Nonoccurrence
Stimulus Problem Sets:		
Frequency of interruptions	89.6	80.8
Percentage of instructions repeated	95.1	88.2
Percentage of instructions completed before beginning	99.8	94.7
Percentage of problems correct	100.0	100.0
Classroom probes:		
Frequency of interruptions	91.6	84.4
Percentage of instructions repeated	94.1	90.2
Percentage of instructions completed before beginning	100.0	100.0
Percentage of problems correct	100.0	100.0

that reliability was acceptable for both high-frequency and low-frequency behaviors. In the first method, agreement was scored only if both judges agreed that the behavior had either occurred or not occurred during a particular problem set. Disagreement was scored only if one judge scored the occurrence of a behavior. Per cent agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements x 100. Agreement was recalculated without counting the number of agreements on nonoccurrence. Therefore, the second set of agreement percentages in Table 3 is based on the total number of agreements plus disagreements.

Classroom probes. One judge rated all presentations of classroom assignments. Two additional judges independently rated one-third of these presentations. The same methods for calculating agreement for all dependent measures as were used on training problem sets were performed. The percentages of agreement for each dependent measure are also summarized in Table 3.

RESULTS AND DISCUSSION

Figures 7, 8, and 9 show the results of the multiple-baseline analysis of the four dependent measures for Subject-2 in baseline, training, and follow-up. While baseline was in effect for problem set-1, a considerable amount of variability was observed for all dependent measures across all problem sets, although a certain degree of stability was noted in some instances (particularly for percentage of problems correct in problem sets 1 and 3). When the training phase was implemented on problem set-1 (Fig. 7.), a rather rapid and considerable improvement in most behaviors (in problem set-1) was observed, and was well maintained across the remaining sessions. An exception was noted for frequency of interruptions, which remained considerably variable (although at a fairly low average of about 1 until training began on problem set-3, when the frequency increased to an average of about 2.75) across all sessions. Generalization effects were also observed across problem sets-2 (a slight decrease in frequency of interruptions, complete and maintained improvement for percentage of instructions repeated, and percentage of instructions completed before beginning, and a somewhat slight increase in percentage of problems correct), and 3 (a rather stable increase in percentage of instructions repeated, and a complete and maintained improvement in percentage of instructions completed before beginning) while training on problem set-1 was in effect.

Figure 7.

Sessions across baseline, training, and follow-up for Subject-2. A multiple-baseline analysis of: frequency of interruptions, percentage of instructions repeated, percentage of instructions completed before beginning, and percentage of problems correct across the first set of problem stimuli.

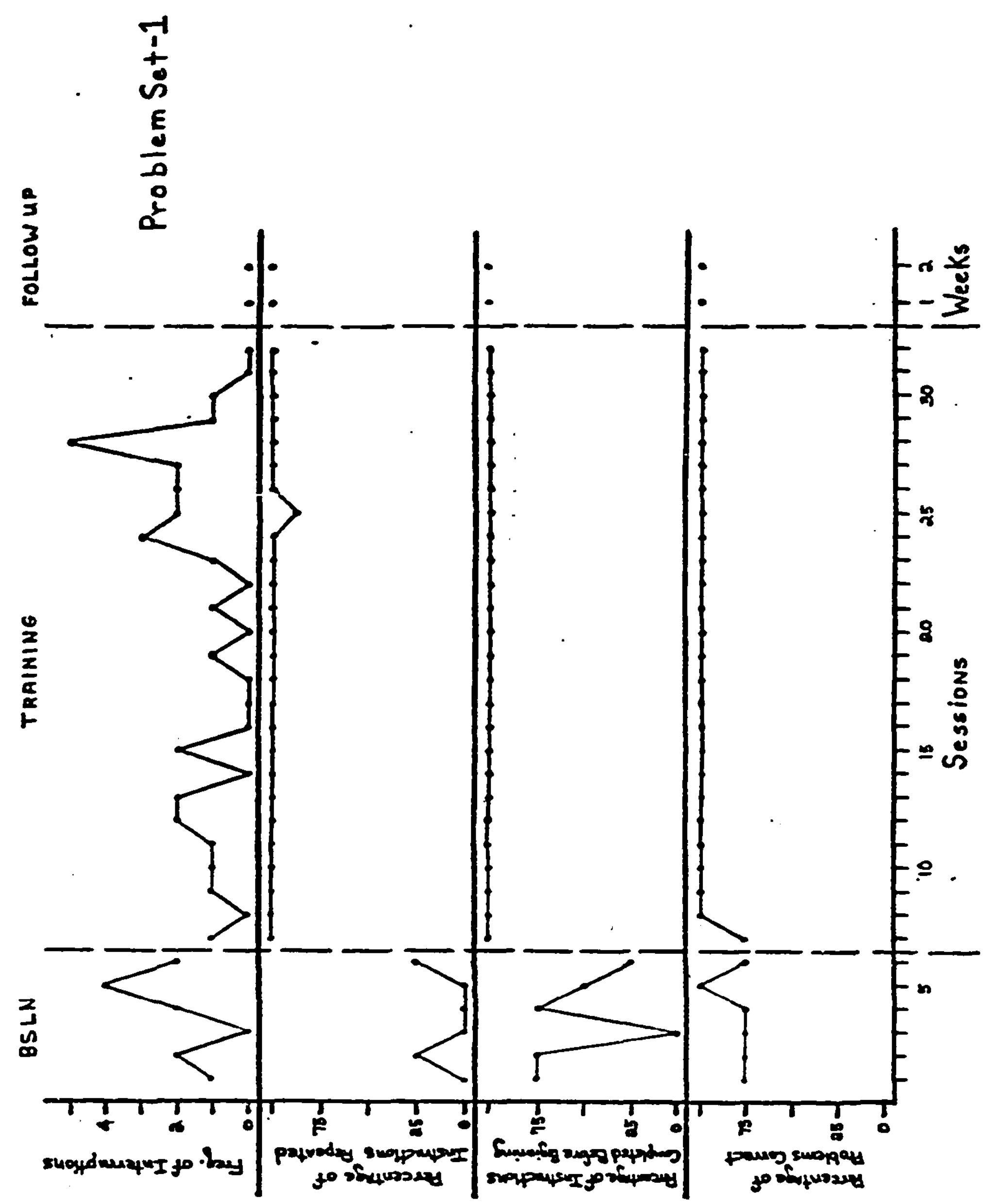
Problem Set-1
BSLN
TRAINING
FOLLOW UP
Freq. of Interruptions
Percentage of Instructions Repeated
Percentage of Instructions Completed Before Beginning
Percentage of Problems Correct
Sessions
Weeks

Fig. 7.

Figure 8.

Sessions across baseline, training, and follow-up for Subject-2: A multiple-baseline analysis of: frequency of interruptions, percentage of instructions repeated, percentage of instructions completed before beginning, and percentage of problems correct across the second set of problem stimuli.

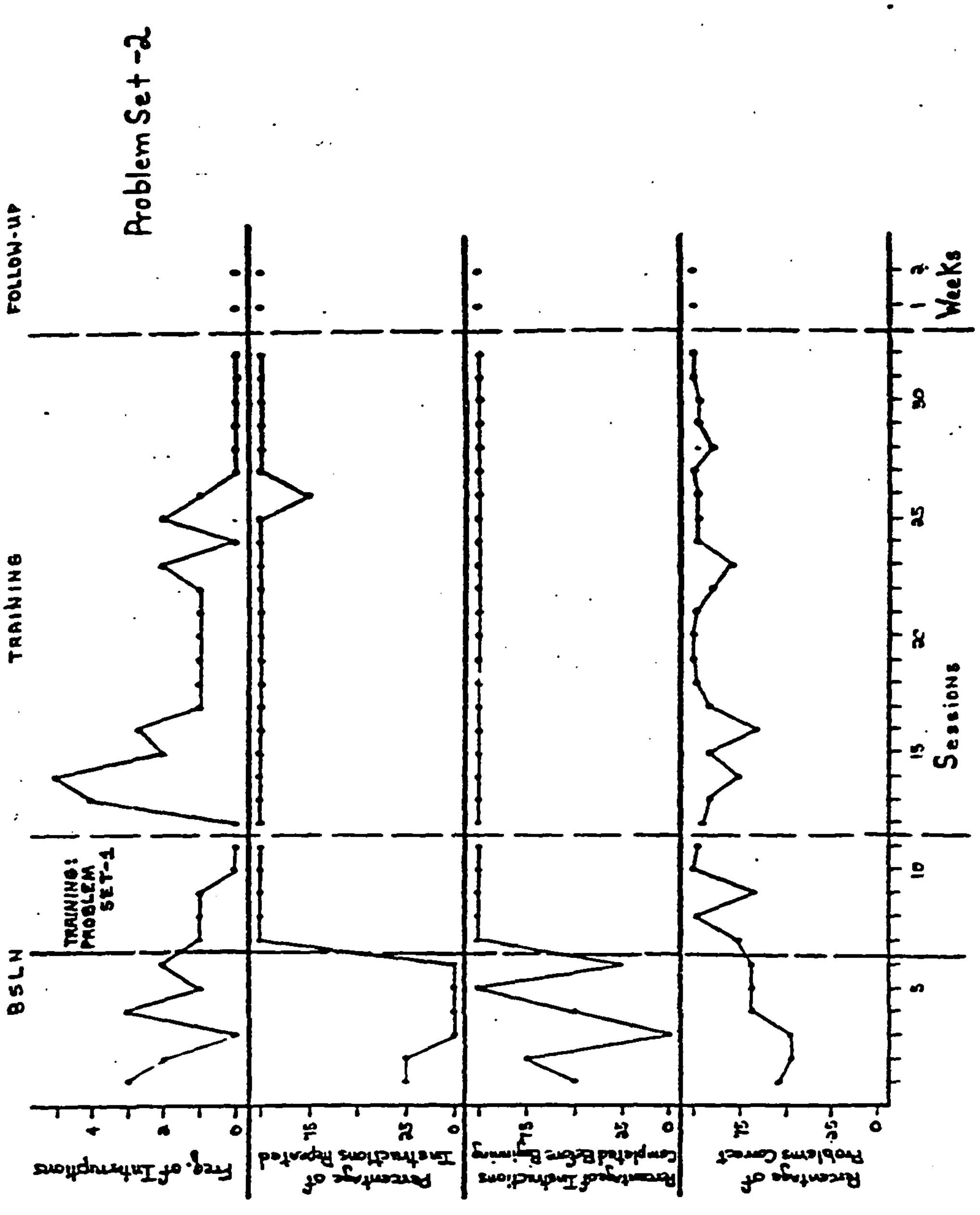
Problem Set -2
BSLN
TRAINING: PROBLEM SET-1
TRAINING
FOLLOW-UP
Freq. of Interruptions
Percentage of Instructions Repeated
Percentage of Instructions Completed Before Beginning
Percentage of Problems Correct
Sessions
Weeks

Fig. 8.

Figure 9.

Sessions across baseline, training, and follow-up for Subject-2. A multiple-baseline analysis of: frequency of interruptions, percentage of instructions repeated, percentage of instructions completed before beginning, and percentage of problems correct across the third set of problem stimuli.

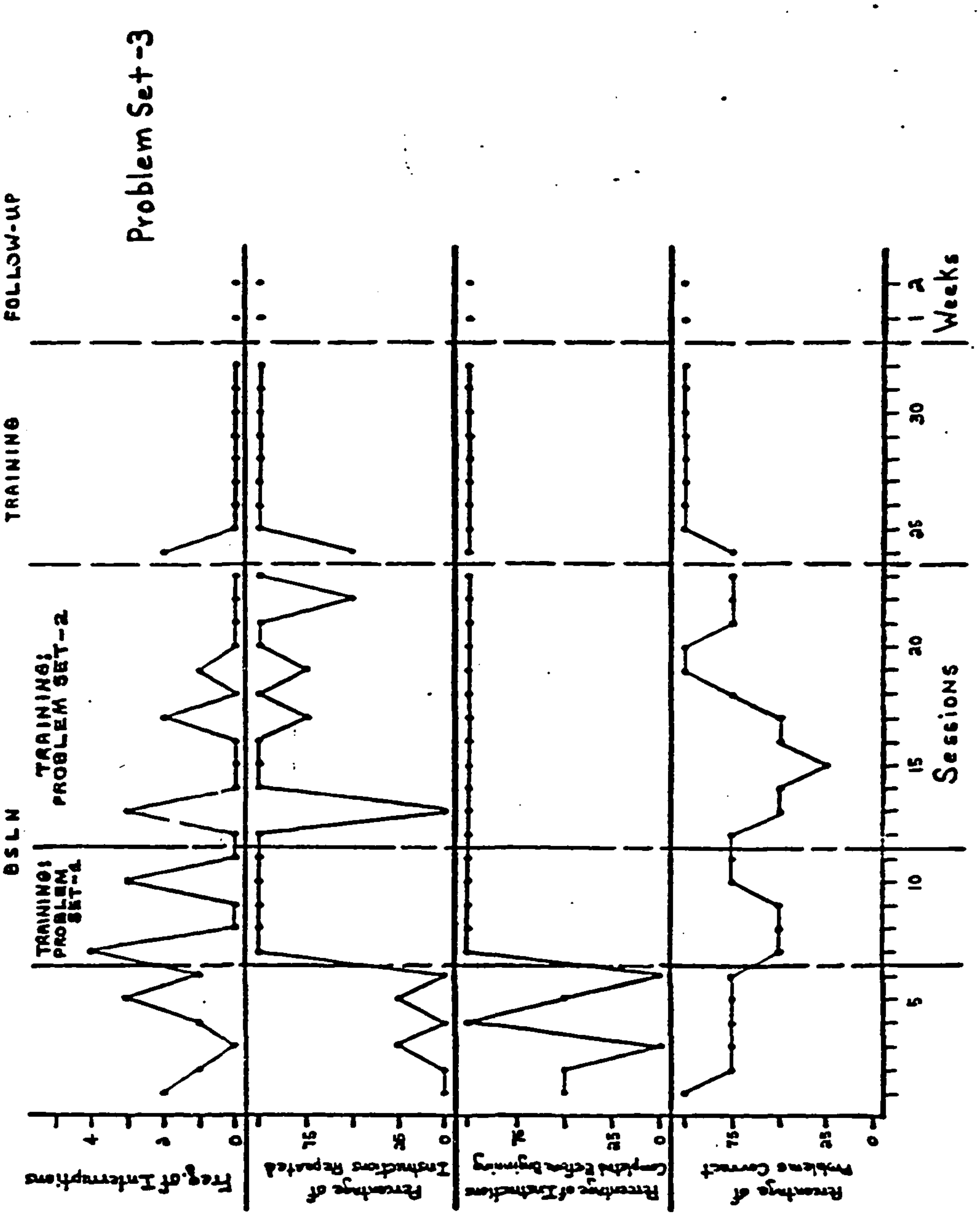
Problem Set-3
TRAINING: PROBLEM SET-1
BSLN
TRAINING: PROBLEM SET-2
TRAINING
FOLLOW-UP
Freq. of Interruptions
Percentage of Instructions Repeated
Percentage of Instructions Completed Before Beginning
Percentage of Problems Correct
Sessions
Weeks

Fig. 9.

When the training phase was implemented for problem set-2 (Fig. 8.), very little change was noted for any behaviors, in terms of improvement, as a result of training on that problem set (a brief, but considerable, increase was noted in frequency of interruptions, however, which gradually declined to an average of about 1 and then remained fairly stable until treatment on problem set-3 began), or in terms of generalization (except for a slight decrease in frequency of interruptions, and a slight increase in percentage of instructions repeated, both on problem set-3).

Upon implementing the training phase for problem set-3 (Fig. 9.), a fairly rapid and maintained improvement occurred for all behaviors (in problem set-3), with generalization effects being apparent with a complete decrease in frequency of interruptions on problem set-2 (however, a considerable increase of frequency of interruptions was simultaneously noted on problem set-1).

Follow-up observations at one and two week intervals indicated that all dependent measures were maintaining completely across all three problem sets.

Table 4 shows the average rate of responses for Subject-2 taken from classroom probe observations (data was collapsed due to the small number of observations: 1 during baseline, 3 during training, and 2 during follow-up).

Table 4

Average percentage rate of responses from classroom probe observations for Subject-2.

Table 4

Dependent Measure	Experimental Phase		
	Baseline	Training	Follow-up
Frequency of interruptions	5.0	1.3	1.5
Percentage of instructions repeated	25.0	67.0	100.0
Percentage of instructions completed before beginning	75.0	83.0	100.0
Percentage of problems correct	33.3	75. 0	87.5

Improvement appears to have occurred for all dependent measures, with the greatest level of improvement noted after training was completed (i.e., during follow-up).

The results of this multiple-baseline analysis indicate that implementation of the training phase coincided with an immediate and maintained improvement in most of the dependent measures, across all three problem sets, with the exception of frequency of interruptions, to some extent in problem set-1, and even less so in problem set-2 (in which an increase was observed). Also, a more gradual improvement was noted for percentage of problems correct in problem set-2. Generalization effects were observed, to some extent, for all dependent measures in problem set-2 after the onset of training in problem set-1, while fewer dependent measures reflected these effects in problem set-3. Fewer generalization effects can be attributed to the onset of training for problem sets-2 and 3, although a slight improvement was noted for frequency of interruptions in problem set-2 during onset of training in problem set-3 (a simultaneous increase in frequency of interruptions was noted for problem set-1, however). A high degree of maintenance of improvement was noted throughout both follow-up observations. Classroom probe data, although based on a small number of observations, indicated a general improvement in all dependent measures, particularly after training ended (i.e., during follow-up). The subject's teacher also gave anecdotal reports which seemed to substantiate that improvements in the subject's responses had been noted in the classroom.

GENERAL DISCUSSION

With the exception of the measure of interruptions for Subject-2 (problem set-1), all dependent measures evidenced considerable improvement, across all stimulus conditions, for both subjects. For the most part, these improvements persisted well into follow-up, although this does not represent maintenance over a very considerable period of time. Classroom probe observations also seemed to indicate that most dependent measures showed a general improvement (particularly during follow-up) over baseline observations for both subjects. These results are somewhat in question however, as both subjects were involved in a number of individualized, classroom behavior programs which focused on improving particular aspects of the subjects' behavior. Subject-2, for example, had an individual behavior program aimed at decreasing interruptions, and more than any other of these programs (for either subject), approximated dependent measures targeted in the present study. After collapsing these data and comparing it to training and classroom probe data, only a very slight resemblance was observed (i.e., individual program data showed the average frequency of interruptions occurring at approximately half the rate as classroom probe data during baseline and training, with average follow-up frequencies being the same for the two measures). Subject-1 was involved in a program whereby he was given feedback (via a checksheet) regarding certain behaviors, including initiating appropriate interactions with others (in which he reportedly did rather well while under close observation). Another program involved being pleasant with staff. Thus, both were only somewhat indirectly related to any of the training stimulus conditions. Both subjects were in the maintenance phase of these individual behavior programs during the course of the present study. There is, nonetheless, a rather good possibility that an interaction existed between the effects of the individual behavior programs and the training the subjects received during this study. Because of this, one can only speculate about the effects that either may have exerted on the observations made in either setting. Also, events particular to Subject-2 (i.e., medication changes, and having his younger brother as a classmate) may have contributed to any of the observed changes in either situation. The extent to which this may or may not be true cannot be readily assessed.

Generalization between dependent measures of trained to untrained stimulus conditions was also observed, although to a greater extent for Subject-2 than for Subject-1. These results can likely be explained in terms of the degree of diversity existing between stimulus conditions presented to the subjects during training and in vivo situations. Whereas Subject-2 was exposed to situations which probably did not sufficiently represent a wide array of typical stimulus occassions, Subject-1 was presented with problems which involved responding to a set of problems which were qualitatively similar in format to most all situations encountered in the classroom.

Hersen, Eisler, and Miller (1974), and Kazdin (1974) found weak generalization effects in stimulus conditions different from those during training. Furthermore, Hersen and Bellack (1976) have suggested that future research should focus on examining the extent to which such training transfers to other, novel (and/or untrained) role-play situations, other role models, and to in vivo (extratherapeutic) behavior. However, such investigations have generally not attempted to examine the degree of diversity which can exist for stimulus training conditions, with an optimum level of generalization occurring to untrained or extratherapeutic stimulus conditions.

Such investigation is necessary, as suggested by Stokes and Baer (1977)...

> Diversity of exemplars seems to be the rule to follow in pursuit of maximum generalization. Sufficient diversity to reflect the dimensions of the desired generalization is a useful tactic. However, diversity may also be our greatest enemy: too much diversity of exemplars of similar responses may make potential gains disproportional to the investment of training effort. (p. 357)

Summarizing, in accordance with a number of studies (Hersen and Bellack, 1976; Frederiksen, et al., 1976; Bornstein, Bellack, and Hersen, 1977), socially (and academically) appropriate responses can be acquired using behavioral rehearsal with instructions, role-reversal and modeling, and performance feedback. Although generalization occurred to untrained situations, the extent to which this occurs, particularly to the classroom (extratherapeutic) setting is questionable, and appears to be relative to the amount of diversity existing between training and extratherapy stimulus conditions.

While preliminary evidence (Foy, et al., 1975) suggests that these learned responses may persist for six months without additional training, additional follow-up data would be necessary to substantiate such possibilities. A clearer separation between training variables, and other setting-related variables (i.e., individual classroom programs, medication changes, etc.) also needs to be achieved.

Finally other questions raised by Stokes and Baer (1977) suggest a direction for future research concerned with training of sufficient exemplars. Namely...

> Is the best procedure to train many examplars with little diversity at the outset, and then to expand the diversity to include dimensions of the desired generalization? Or is it a more productive endeavor to train fewer exemplars that represent a greater diversity? (p. 357)

BIBLIOGRAPHY

Allen, G. J. Case study: Implementation of behavior modification techniques in summer camp settings. Behavior Therapy, 1973, 4, 570-575.

Baer, D. M., Peterson, R. F., and Sherman, J. A. The development of imitation by reinforcing behavioral similarity to a model. Journal of the Experimental Analysis of Behavior, 1967, 10, 405-416.

Bailey, J. S. A handbook of research methods in applied behavior analysis, New York: Plenum Publishing Co., 1977.

Bornstein, M. R., Bellack, A. S., and Hersen, M. Social-skills training for unassertive children. Journal of Applied Behavior Analysis, 1977, 10, 183-195.

Bucher, B., and Mueller, H. H. Acquisition and generalization of compliance with complex five-component instructions. Journal of Applied Behavior Analysis, 1977, 10, 507.

Clark, H. B., and Sherman, J. A. Teaching generative use of sentence answers to three forms of questions. Journal of Applied Behavior Analysis, 1975, 8, 321-330.

DuNann, D. H., and Weber, S. J. Short and long-term effects of contingency managed instructions, on low, medium, and high GPA students. Journal of Applied Behavior Analysis, 1976, 9, 375-376.

Foy, D. W., Eisler, R. M., and Pinkston, S. G. Modeled assertion in a case of explosive rages. Journal of Behavior Therapy and Experimental Psychiatry, 1975, 6, 135-137.

Frederiksen, L.W., Jenkins, J. O., Foy, D. W., and Eisler, R. M. Social-skills training to modify abusive verbal outbursts in adults. Journal of Applied Behavior Analysis, 1976, 9, 117-125.

Frisch, S. A., and Schumaker, J. B. Training generalized prepositions in retarded children. Journal of Applied Behavior Analysis, 1974, 7, 611-621.

Garcia, E. The training and generalization of conversational speech form in non-verbal retardates. Journal of Applied Behavior Analysis, 1974, 7, 137-149.

Garcia, E., Baer, D.M., and Firestone, I. The development of generalized imitation within topographically determined boundaries. Journal of Applied Behavior Analysis, 1971, 4, 101-112.

Griffiths, H. and Craighead, W.E. Generalization in operant speech therapy for misarticulation. Journal of Speech and Hearing Disorders, 1972, 37, 485-494.

Hersen, M., and Bellack, A.S. A multiple-baseline analysis of social-skills training in chronic schizophrenics. Journal of Applied Behavior Analysis, 1976, 9, 239-245.

Hersen, M., Eisler, R.M., and Miller, P.M. An experimental analysis of generalization in assertive training. Behavior Research and Therapy, 1974, 12, 295-310.

Horton, G.O. Generalization of teacher behavior as a function of subject matter specific discrimination training. Journal of Applied Behavior Analysis, 1975, 8, 311-319.

Johnston, J.M., and Pennypacker, H.M. Strategies and Tactics of Human Behavioral Research. 1979 (in press).

Kazdin, A.E. Effects of covert modeling and model reinforcement on assertive behavior. Journal of Abnormal Psychology, 1974, 83, 240-252.

Koegel, R.L., and Rincover, A. Setting generality and stimulus control in autistic children. Journal of Applied Behavior Analysis, 1975, 8, 235-246.

Murdock, J.Y., Garcia, E.E., and Hardman, M.L. Generalizing articulation training with trainable mentally retarded subjects. Journal of Applied Behavior Analysis, 1977, 10, 717-733.

Rimm, D.C., and Masters, J.C. Behavior Therapy: Techniques and Empirical Findings. New York: The Academic Press, 1973.

Skinner, B.F. Science and Human Behavior. New York: The Free Press, 1953.

Stokes, T.E., and Baer, D.M. An implicit technology of generalization, Journal of Applied Behavior Analysis, 1977, 10, 349-367.

Stokes, T.E., Baer, D.M., and Jackson, R.L. Programming the generalization of a greeting response in four retarded children, Journal of Applied Behavior Analysis, 1974, 7, 599-610.

Tracey, D. A., Briddell, D. W., and Wilson, T. G. Generalization of verbal conditioning to verbal and nonverbal behavior: Group therapy with chronic psychiatric patients. Journal of Applied Behavior Analysis, 1974, 7, 391-402.

Wahler, R. G. Some structural aspects of deviant child behavior. Journal of Applied Behavior Analysis, 1975, 8, 27-42.