THE UNIVERSITY OF MANITOBA

MEDIATE GENERALIZATION, TRAIN SUFFICIENT EXEMPLARS, AND TRAIN-TO-GENERALIZE AS STRATEGIES FOR PROGRAMMING GENERALIZATION IN THE IMPROVISATION OF TOOLS BY PRESCHOOL CHILDREN

by

Glenn B.B. Allard

A Thesis

Submitted To The Faculty Of Graduate Studies In Partial Fulfillment

Of The Requirements For The Degree Of Master Of Arts.

Department of Psychology

Winnipeg, Manitoba

December, 1979

MEDIATE GENERALIZATION, TRAIN SUFFICIENT EXEMPLARS, AND TRAIN-TO-GENERALIZE AS STRATEGIES FOR PROGRAMMING GENERALIZATION IN THE IMPROVISATION OF TOOLS BY PRESCHOOL CHILDREN

BY

GLENN BRIAN BERNARD ALLARD

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

©1980

Permission has been granted to the LIBRARY OF THE UNIVERSITY OF MANITOBA to lend or sell copies of this thesis, to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film, and UNIVERSITY MICROFILMS to publish an abstract of this thesis.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

ACKNOWLEDGEMENTS

I would like to thank the staff at the Greta Brown Unit day-care centre, for allowing me to work with the children there and for providing enthusiastic encouragement and interest in my work. While there I was able to see the value and important role that good day-care programs have for children today.

I am grateful to my examination committee; Drs. Steve Holborn, Joe Pear and Larry Cooley, for their helpful suggestions. I would especially like to acknowledge two members of my committee. First, the painstaking and meticulous work of Dr. Holborn on this manuscript and his assistance in solidifying my conceptual arguments were much appreciated. Second, Dr. Cooley's encouragement, friendship and sharing of the many philosophical questionings I face was invlauable to me.

Finally I would like to express my appreciation to Joanne Swain for her many contirbutions to me and my work. Through her assistance in data collection and idea development she helped me to complete this thesis. More importantly, however, her constant encouragement and unconditional love during the many months of work gave me a significant purpose in completing this thesis.

TABLE OF CONTENTS

Page

Acknowledgements i

Table of Contents ii

Abstract v

List of Tables viii

List of Figures ix

A Behavior Analysis of Problem Solving 1

Training Creativity and Problem Solving 6

Story Writing 8

Block-Building 10

Painting 11

Generalization Programming 13

Program Common Stimuli 14

Train Sufficient Exemplars 14

Train-To-Generalize 16

Mediate Generalization 16

Training Generalized Problem-Solving 17

Generalized Improvisation of Tools 18

Discrimination Training 27

Generalized Improvisation and Discrimination Behavior 29

Hypotheses 30

Method 34

Subjects and Setting 34

Equipment and Performance Criteria 36

Improvisation Tasks 39

Hammers 39

Containers 39

Page

Behavioral Definitions 39

Improvisations 40

Distractors 40

Exemplars 40

Data Collection Procedures 40

Inter-Observer Reliability 41

Training Procedure 42

Session Introduction 42

Baseline 43

Improvisation Training 46

Train-To-Generalize 46

Train Sufficient Exemplars 47

Sufficient Exemplars Plus Descriptive Praise 47

Mediate Generalization 48

Reinforcement Only 49

Multiple Baseline Reversal Design 50

Results 52

Dependent Measures 52

Sufficient Generalization Criterion 53

Multiple Baseline Across Tasks Data 56

Train-To-Generalize 56

Jason 56

Michelle 59

Train Sufficient Exemplars 61

Sylvia 61

Aysa 63

Sufficient Exemplars Plus Descriptive Praise 66

Wendy 66

Che 68

Daniela 70

Mediate Generalization 72

Delia 72

Alvin 74

Tanya 76

Multiple Baseline Across Subjects Data 78

Train-To-Generalize 78

Train sufficient Exemplars 81

Sufficient Exemplars Plus Descriptive Praise 84

Mediate Generalization 86

Strategy Comparison Data 87

Discussion 93

Train-To-Generalize 93

Train Sufficient Exemplars 94

Sufficient Exemplars Plus Descriptive Praise 96

Mediate Generalization 98

Problem Situations and Problem Solving 100

Discrimination Training and Generalization Programming 101

Methodology and Future Research 102

Conclusion 105

Reference Notes 107

References 107

Appendix A 110

Appendix B 115

Abstract

The development of new and complex creative behaviors via behavior modification techniques has been demonstrated in children's story writing, painting, block building and tool improvisation. From problem solving and generalization perspectives one of the more interesting classes of behaviors trained is tool improvisation, which is defined as finding a new and effective alternative tool in the absence of the original tool. Viewed behaviorally, problem solving is an operant and as such is amenable to development through differential reinforcement. Creative problem solving, as in tool improvisation, can be viewed as the generalization of original and functional problem-solving responses within a task and across other tasks. The present study trained 10 preschool children to produce numerous, new alternative tools in two tasks, one using functional hammers and one using functional containers. Improvisation was studied in two multiple baseline designs, one across tasks for each child and one across subjects for each task. Four generalization programming strategies were studied in the multiple baseline design. The children were randomly assigned to one of the following strategies: train-to-generalize; train sufficient exemplars; sufficient exemplars plus descriptive praise; and mediate

generalization. The first strategy involved socially reinforcing the child in a way descriptive of the child's response when a new tool was produced. The second strategy involved demonstrating new alternative responses with no descriptive social reinforcement. The third strategy used both examples of alternative tools and descriptive social reinforcement. The fourth strategy used examples of alternative tools with verbal labels of the item's essential functional quality (e.g., hardness for hammers and spaciousness for containers) in addition to descriptive social reinforcement. The dependent measures were: cumulative number of improvisations; percent successful discrimination; and mean rate of improvisation per session during training. The measure of successful discrimination, taken in conjunction with cumulative improvisation, was of significant value in establishing the effectiveness of training and in defining a criterion of sufficient generalization for each task. The results revealed that with a very low operant level train-to-generalize alone was virtually ineffective in producing substantial tool improvisation. Train sufficient exemplars was more successful than expected as was sufficient exemplars plus descriptive praise. The success of these latter two strategies appeared to be due to modeling of exemplar responses which then elicited child responses. Mediate generalization did not produce substantially more generalization within the tasks than train sufficient exemplars and sufficient exemplars plus descriptive praise. However, the use of verbal labels in

mediate generalization demonstrated the most promise for establishing generalization across tasks. The potential of mediate generalization as a powerful generalization programming strategy was discussed in relation to its success in generalization across tasks as well as its development of verbal knowledge about the problem solving process. All strategies produced much greater improvisation than achieved in Parsonson and Baer's (1978) original attempt to establish tool improvisation in preschool children. The success of the present strategies was mainly attributable to the modeling of example responses (picking exemplar items from the same tray as provided the child) and the structure of the problem situation (removing solution items to maintain a new problem situation).

LIST OF TABLES

Page

Table 1. Examples of materials used: exemplars; potential tools; and distractors in two tool tasks 38

Table 2. Experimental conditions for each subject 44

Table 3. Summary of frequency of new improvisations in training . . . 90

Table 4. Summary of frequency of new improvisations in training, subjects in Parsonson and Baer (1978) study 92

LIST OF FIGURES

Page

Figure 1. Cumulative number of improvisations and percent successful discrimination for Jason in hammers and containers . . 57

Figure 2. Cumulative number of improvisations and percent successful discrimination for Michelle in hammers and containers 60

Figure 3. Cumulative number of improvisations and percent successful discrimination for Sylvia in hammers and containers .. 62

Figure 4. Cumulative number of improvisations and percent successful discrimination for Aysa in hammers and containers . . 64

Figure 5. Cumulative number of improvisations and percent successful discrimination for Wendy in hammers and containers . . 67

Figure 6. Cumulative number of improvisations and percent successful discrimination for Che in hammers and containers . . . 69

Figure 7. Cumulative number of improvisations and percent successful discrimination for Daniela in hammers and containers . 71

Figure 8. Cumulative number of improvisations and percent successful discrimination for Delia in hammers and containers . . 73

Figure 9. Cumulative number of improvisations and percent successful discrimination for Alvin in hammers and containers . . 75

Figure 10. Cumulative number of improvisations and percent successful discrimination for Tanya in hammers and containers . . 77

Figure 11. Cumulative number of improvisations and percent successful discrimination for Train-To-Generalize in hammers . . 79

Figure 12. Cumulative number of improvisations and percent successful discrimination for Train-To-Generalize in containers . 80

Figure 13. Cumulative number of improvisations and percent successful discrimination for Train Sufficient Exemplars in hammers 82

Page

Figure 14. Cumulative number of improvisations and percent successful discrimination for Train Sufficient Exemplars in containers 83

Figure 15. Cumulative number of improvisations and percent successful discrimination for Sufficient Exemplars and Descriptive Praise in hammers 85

Figure 16. Cumulative number of improvisations and percent successful discrimination for Sufficient Exemplars and Descriptive Praise in containers 86

Figure 17. Cumulative number of improvisations and percent successful discrimination for Mediate Generalization in hammers 88

Figure 18. Cumulative number of improvisations and percent successful discrimination for Mediate Generalization in containers 89

The ability to solve old and new problems efficiently and in different ways is a highly valued skill in our society, that has received much attention and experimental study since 1950 when J.P. Guilford indicated the importance of creativity to future social development (Taylor, 1975). Behavior modification techniques have only been applied in a small number of creativity studies (Goetz, Note 1) and have been criticized in some quarters, for attempting to control a process which requires freedom (Cole & Parsons, 1974). But if creativity is a desired goal for individuals and society, then it seems only reasonable that it should be increased whether that is "control" or not. In fact, despite the small number of studies and the philosophical concern over control, the behavior analyses and training conducted have clearly demonstrated that behavior modification is a successful strategy for increasing creativity, and as such deserves further intense examination to broaden its applications and findings (Goetz, Note 1).

A Behavior Analysis of Problem Solving

Skinner (1953, p. 247) defined problem solving as, "any behavior which, through the manipulation of variables, makes the appearance of a solution more probable." The definition of problem solving becomes clearer in the context of what a problem is. A problem is said to exist when a response can-

not be emitted and the state of deprivation or aversive stimulation is strong. The problem is solved by changing the situation to allow the response to occur, which then alters the deprivation or aversive stimulation (Skinner, 1953; Bijou, 1976). Consider the following example of a problem from Skinner (1953, pp. 246-247):

> A locked drawer presents a problem if behavior requiring an open drawer is strong and if the individual does not have the key or other means of opening it. The strength of the behavior is inferred from the presence of responses which previously opened the drawer or from the appearance of the behavior as soon as the drawer has been opened... In any case, the solution to a problem is simply a response which alters the situation so that the strong response can be emitted... Once the solution has occurred, the problem vanishes simply because the essential condition has been eliminated. (The same problem is not likely to recur since the situation will no longer be novel. Henceforth, the response which appeared as a solution will occur because it has been reinforced under similar circumstances).

There are three important points to be noted in Skinner's example of a problem. First, the example indicates that the subject has a response in repertoire which cannot be emitted, in this case getting into the drawer to acquire the reinforce-

ment within it. The subject must have facility at getting into the drawer and must have been previously reinforced for doing so, or the state of deprivation will not exist and so the problem does not exist. The second point is that there is a distinction between the strong response which is getting into the drawer, and the solution response which is the alteration of the situation so that the key is found or some other means of opening the drawer is made available. That is, the final response of getting into the drawer is considered the strong response not the solution. The solution response is that which occurs before the strong response (getting into the drawer) i.e., finding the key or some other means which allows the drawer to be opened. When studying problem solving then, the concern should be with the process; that is, the step(s) taken to alter the situation so that the strong response can occur. The third and final point is the essential condition of a problem situation, which is, that no solution must be present. Any previously reinforced solutions are likely to occur, (e.g., using a paper clip as a key), therefore to achieve a new problem solution the problem situation must be new in some way. If a previously performed solution is possible, then, in fact, no problem exists.

The appearance of a solved problem (e.g., an open drawer) does not verify that a systematic problem solving process has occurred, because the solved problem may have resulted accidentally from random responding. Accidentally solved problems are

the result of trial-and-error responding (Skinner, 1953; 1968; Bijou, 1976). Because a problem situation involves a state of deprivation or aversive stimulation, the subject will be responding to eliminate the deprivation or aversive stimulation. The result may be that a randomly produced response or series of responses may serve to dispose of the essential condition of the problem. The problem has been solved, but only accidentally and inefficiently. All that can really be said is that the subject has learned to try, until a solution appears. Any further refinements of the process, which inevitably would be necessary, will be time consuming and frustrating to the subject, because previously reinforced non-essential responses must be extinguished.

To be properly called problem solving, a response or a series of responses must systematically change the environment to allow the solution response to occur. That is, there should be a minimum of inappropriate attempt responses. In trial-and error responding where a solution occurs, many inappropriate and unnecessary responses have been adventitiously reinforced. The series of responses learned will not facilitate the solving of other problems, rather the subject, when faced with another problem, will again randomly try responses until a solution appears.

Skinner (1953; 1968) has indicated that "problem solving" refers to three categories of precurrent activities (i.e., activites before the strong response) which facilitate problem

solving under a variety of circumstances. The three categories of precurrent activities are: (a) surveying the problem situation, (b) arranging or re-arranging the situation, and (c) self-probing (i.e., systematically reviewing tentative solutions). A simple example of a precurrent response is attention, which is included in the first category above, surveying the problem situation. What is meant specifically by attention is that a subject must recognize the kind of problem involved, then must respond to selected features of the environment. Proper attending can be learned. Usual methods of teaching it involve showing an example and gaining attention by talking loudly or requiring that the subject look toward the example. These methods of getting someone to attend do not guarantee that the subject has attended to the problem or to the step(s) by which a solution was reached.

Skinner (1968) considers a more fruitful approach to be one of specifying discrimination as the central process at work in attending. That is, the subject must learn to discriminate between essential and non-essential features of the situation. One way to facilitate discrimination in attending, suggested by Skinner (1968), is to use instructions. Specific instructions on how to attend and to what, are merely arranging present contingencies to take advantage of similar complex contingencies of stimulus control in the subjects past history. Problem solving is thus facilitated because the instructions form supplemental stimuli in the present and future problem situa-

tions, which aid in eliminating the essential condition of the problem.

It is not readily obvious to what features of particular situations attention should be trained and thereby problem solving facilitated. However, the first step in the behavioral analysis, in terms of the defining characteristics of precurrent activities, is to consider problem solving as operant behavior. That is, in the definition of precurrent activities the behavior in question must alter the environment in some way, which corresponds to the definition of operant behavior (Ferster, Culbertson & Perrott-Boren, 1975). An operant behavior may take on any form, and it conceivably may be desirable to train as many forms as possible to facilitate a wide range of problem solving, in many circumstances (Skinner, 1953). The training of diverse responding in creative tasks and problem solving contexts has followed from this analysis.

Training Creativity and Problem Solving

As just indicated, the training of creative behavior and problem solving has followed from defining them as operant behaviors. Differential reinforcement of the actual behavior has been the major training strategy employed, with some use of exemplar training, but no direct instructions on attending appear to have been formulated.

In the study and training of creativity it is important to

clearly define creativity. Bijou (1976) has indicated that creative behavior is a special case of problem solving. Creative behavior is a term applied to a solution to a problem, that is both original and functional. Most creative behavior and problem solving involves a complex series of responses. However, the simplest definition of creativity entails new responses, relative to a subject's past responses, which are systematically (not trial-and-error) produced (Goetz, Note 1).

To determine whether a solution for a given problem is unique to a person it is necessary, then, to have observed all of that person's problem solving episodes over some period of time, which is typical of behavior modification research designs. Traditionally, however, creativity has been viewed as an ill-defined cognitive quality assessed by means of accounts of personal experience or standardized tests (Goetz, Note 1). The difficulty with such views is that a score on a particular creativity test at a particular time means little in terms of external validity, i.e., an individual actually being able to make new things or to act in new ways (Torrance, 1969).

From earlier discussion we can see that cognitive abilities are analyzable in terms of their component behaviors and that to make progress in training creativity, behavior and its products must be the targets of training (Goetz, Note 1; Holman, Goetz & Baer, 1977). In particular, the products of behavior readily lend themselves to objective measurement in terms of the analysis of new, unique forms and functionality.

This kind of objective measurement is important because not all idiosyncratic novelty is considered socially valuable (Skinner, 1968; Guilford, 1975). Using products to measure creativity within behavior modification research designs has proved prosperous to the training of story writing, block-building and painting with children.

Story Writing. Composition writing in elementary school children has been analysed in terms of its component parts and these parts have been taught sequentially using differential reinforcement. The first training of story writing involving behavior analysis appears to have been undertaken by Bringham, Graubard and Stans (1972). The composition components identified and then taught sequentially were number of words used; number of different words used; and number of new words used. Token reinforcement, as team points, was placed on the three composition components sequentially in time and the results showed that in turn the usage of each component by the children increased. This study demonstrated the benefit of a rigorous analysis of a complex behavior considered to involve being creative.

Also in composition writing Maloney and Hopkins (1973) meticulously identified eight components in children's written compositions. They measured: the number of letters; the number of words; the number of adjectives; the number of adverbs; the number of action verbs; the number of prepositional

phrases; the number of sentences; and the number of compound sentences. Again team points were used to sequentially increase the usage of each component. Maloney and Hopkins added significantly to the demonstration that behavior modification could increase individual creativity, by having the stories written, independently rated subjectively for creativity. The subjective ratings indicated that as more components were taught, the stories were judged as more creative relative to earlier stories written by that individual.

Next Ballard and Glynn (1975) attempted to determine exactly what components produced more creative stories and if self-control strategies could contribute to creative composition writing. Their response measures were: the number of sentences; the number of different action words; and the number of different describing words. They used self-determined and self-administered reinforcement, self-assessment and self-recording as the interventions to increase the usage of the composition components. Their results showed that only when reinforcement was applied did usage of the components increase. Also they found that stories written with reinforcement contingent on the number of different action words received the highest qualitative ratings.

The story writing studies have shown that complex human behaviors can be broken down into their essential component parts and that these parts can be taught sequentially using reinforcement. Increased usage of some or all of the components

produced more creative stories as indexed by subjective ratings. Other areas of children's creative behavior have been similarly analysed and trained using reinforcement.

Block-building. Goetz and Baer (1973) demonstrated how descriptive social reinforcement could increase form diversity and the number of forms used in children's block-buildings. Descriptive social reinforcement, i.e., describing the behavior that is being reinforced (e.g. "Good girl, you made a new tower by putting that block on top of those two blocks!") was employed. This type of social reinforcement comes very close to Skinner's (1968) suggestion of using instructions to specify the essential features of the behavior desired.

Form diversity was defined as the use of two or more blocks to create a specified shape or function. They listed twenty possible forms, then provided descriptive social reinforcement on the occurrence of each form in each session. The results indicated that contingent descriptive social reinforcement increased form diversity within sessions and increased the total number of forms used across sessions. They argued against a definition of creativity which compared an individual to a group norm. Novelty defined as something never seen before in a particular society, may lead to very infrequent reinforcement for most individuals. Rather novelty as something not seen in an individual's repertoire in each new situation will lead to more frequent reinforcement and increase the

probability of future creative behavior.

Painting. A number of experiments in children's painting have been conducted by Holman et. al. (1977). In one experiment they reinforced form diversity in easel painting with descriptive social reinforcement, while monitoring the effect on block-building which received no reinforcement. Form diversity and number of new forms increased in easel painting as training went on; however, only a slight increase in block-building form diversity was noted. The concern was then raised that behavioral training of creativity needed to produce more generalization to untrained behaviors, to be properly recognized by society as creativity.

A second experiment was carried out to assess if generalization of creative responding would occur if two tasks were similar enough in topography. Descriptive social reinforcement was given to form diversity in felt-tip pen drawing while monitoring easel painting, block-building and lego construction. The manipulation demonstrated increased form diversity in felt-tip pen drawing, and somewhat variable increases in the other three tasks. The most consistent generalization increases were noted in easel painting, a topographically similar response to felt-tip pen drawing.

The major contributions of these two studies were; their examination of generalization effects in behavioral training of creativity; and their suggestion to consolidate research

by fitting creativity training strategies into the Stokes and Baer (1977) review of generalization programming. Holman et.al. also encouraged the notion that creativity training should work in collaboration with generalization programming and that the training had to overcome limitations of topograhical similarity of the behaviors trained, while emphasizing functional similarity. They questioned the social value of any creativity training which was limited to one kind of task or where generalization occurred only in topographically similar tasks; like felt-tip pen drawing and easel painting. This questioning seems somewhat unwarranted, however, because of what is known about stimulus control and differential reinforcement, i.e., that responses occur in the presence of certain stimuli and not others because they have been previously differentially reinforced for doing so. In fact it is usually undesirable to obtain generalization to all situations, e.g., waving good-bye at inappropriate times. Bijou (1976) clearly indicated in his analysis of creativity and problem solving that since creative behavior is not viewed as an expression of a creative faculty, but rather as the behavior of an individual in relation to specific environmental events, then all references to a child's problem solving ability must be in specific terms. It is not meaningful to say a child is creative, we must specify the area of creative behavior - in music, painting, science, mechanical technology, and so on.

Indeed there is social value and it is reinforcing for the individual if creativity is developed in only one area. For example someone who can produce great oil paintings does not need to also be a great musician for substantial individual and social benefits to accrue.

The problems in obtaining generalization in creative responding bring up two important issues in the training of creative behavior through differential reinforcement and/or stimulus control procedures (Goetz, Note 1). The first issue of concern is that the operant level of the response desired may be so low that it can never be reinforced. The second issue is that creativity training requires the teaching of an abstraction rather than a specific set of behaviors. That is, the subject must discriminate between the novelty of the response (which is being reinforced) and the response itself. Both Holman et. al. (1977) and Goetz (Note 1) see the solutions to these issues coming through generalization programming.

Generalization Programming

Stokes and Baer (1977) reviewed the generalization programming strategies used in previous research and then classified them according to similarity of procedures. Four of the strategies reviewed seem especially pertinent to creativity research in the light of studies already completed and their failure to obtain generalization. These four strategies will be summarized here.

Program Common Stimuli. If there are sufficient stimulus components occurring in common in both the training and generalization settings, then generalization will be enhanced (Stokes & Baer, 1977). Very few behavioral studies have deliberately attempted to program common stimuli in training and generalization settings; especially common physical stimuli (Stokes & Baer, 1977). If these stimuli were in some way incorporated into the training procedures, then generalization might thereby be facilitated.

Train Sufficient Exemplars. Another generalization strategy outlined in Stokes and Baer (1977) was the training of a sufficient number of examples (exemplars) of a desired response class. That is, a lesson is taught; if no generalization occurs, then another lesson is taught until there is sufficient generalization. There are two types of exemplars: stimulus and response. Stimulus exemplars are used to obtain a particular response in the presence of a variety of stimuli, for example, getting a person to wave hello to all the staff in an institution, not just one person. Response exemplars are used to generate a variety of responses similar to, but not necessarily the same as, an example response in a given situation. For example a child might be taught the correct use of plurals for a set of nouns and then generalize to all nouns. In both types of exemplars the best ordering and combination of the presentation of exemplars has not been clearly researched.

A significant point which requires clarification is what constitutes a sufficient number of exemplars; and what constitutes a sufficient amount of generalization. Stokes and Baer (1977) said that the number of exemplars is a few, usually less than three and that the extent of exemplar training may depend on the subjects experience with the task and the nature of the task itself. In two studies cited by Stokes and Baer (1977) as examples of response exemplar training (i.e., Baer, Peterson & Sherman, 1967; Guess, Sailor, Rutherford & Baer, 1968) sufficient generalization lay between 80 and 100 percent correct responding and the number of response exemplars varied greatly. Baer et.al. (1967) trained imitation responses in retardates and found that generalized imitation to unreinforced responses began after approximately twenty responses had been trained. Guess et.al. (1968) trained the correct use of plurals in a retarded girl. They found that after only three examples were trained, the girl correctly used singular and plural nouns without further training. Most recently Parsonson and Baer (1978) used the train sufficient exemplars technique but provided no index of sufficient generalization or of the sufficient number of exemplars required. The establishment of an index of sufficient generalization and thereby of sufficient exemplars seems crucial to the development of the technique.

Train-to-Generalize. Stokes and Baer (1977) defined this strategy as considering generalization as a response which can be increased through differential reinforcement. Application of the strategy was exemplified by Goetz and Baer (1973). Basically the application of train-to-generalize involved offering descriptive social reinforcement only for new responses made by the child on the task at hand. The child is rewarded for moving along the generalization gradient underlying block form inventions, and never for staying at any one point. Holman et.al. (1977) recommended the procedure as possibly the most productive in training creative responding. However, a serious problem would seem to be that the operant level of the desired response may not be high enough to be affected by the positive reinforcement.

Mediate Generalization. This procedure requires establishing a response as part of the new learning that will likely be utilized in other problem settings as well. Generalization will result if the response constitutes sufficient commonality between the original learning and new problem situations (Stokes & Baer, 1977).

The most commonly used mediator is language, although the deliberate application of language is rare. Verbal mediation should not have tightly restricted discriminative value for the subject or it will fail. The verbal response must be salient to both the training and the generalization problems (Stokes

& Baer, 1977). Language is a response, but it is also a stimulus both for the speaker and the listener; therefore it meets perfectly with the logic of a salient common stimulus to be carried by the subject from any training setting to any generalization setting.

The mediate generalization strategy has not been used in creativity training. However, it seems emminently suitable because in addition to the above reasons, conceptually it follows Skinner's (1968) suggestion of constructing supplemental stimuli in the form of instructions to facilitate problem solving.

Training Generalized Problem-Solving

Before problem-solving can occur or be trained the subject must have the various behaviors that lead to the solution in repertoire (Bijou, 1976) or they must be shaped. Specifically, precurrent activities (Skinner, 1953) altering the environment to allow a solution to appear, might involve one or more of the following manipulations (Bijou, 1976): (a) moving physical objects or physical events (e.g., manipulating materials which might be covering the key to a locked drawer, or grasping the drawer handles and pulling to open the drawer); (b) arranging social conditions (only telling bad news in certain ways or in certain places); (c) altering personal conditions (not eating snacks in order to lose weight); or (d) manipulating abstract conditions (transposing numbers to solve arithmetic

problems). If a child's learning history does not include situations where the required responses were developed then the problem is insoluble to the child until those responses are developed. A child's effectiveness in problem solving then depends on skills developed which are appropriate to the type of problem faced or similar problems, and on the reinforcement history for solving past problems. The development of these skills will affect the success of training as was noted by Stokes and Baer (1977) in discussing the train sufficient exemplars technique (cf., p. 15 of present study).

One class of behaviors which constitutes an important part of a problem-solving repertoire with physical objects, is improvisation. The development of improvisation in preschool children was studied by Parsonson and Baer (1978).

Generalized Improvisation of Tools

Using a combination of two generalization strategies (train-to-generalize and train sufficient exemplars) Parsonson and Baer (1978) trained problem solving in preschool children as the improvisation of tools. Improvisation was defined as producing a functional alternative tool from a tray of items in the absence of the regular tool. They stated: "The process of improvisation involves identifying the essential characteristics of the unavailable item, and then searching for the effective alternative." (Parsonson & Baer, 1978, p. 364).

Five preschool children were involved to determine whether they spontaneously displayed diverse improvisation skills. If not, it could then be determined whether they could be trained to display diverse and new improvisations and if so, whether generalized improvisation in one problem task would generalize its effects to untrained tasks. They used three tasks; pounding on a pounding toy; storing marbles; and lacing a shoe. Each task required the improvisation of that task's tool, that is, of hammers, containers and shoelaces.

There were two experimental designs used in combination: a multiple baseline across tools for each of the five children; and a multiple baseline across two subjects for each of the tool tasks. In each baseline session the experimenter demonstrated the use of the regular tool (baseline exemplar), then invited the child to try using the tool. Successful completion of the task was accompanied by descriptive feedback and praise, (e.g., "Good girl! You have hit the peg right down with the hammer."). An improvisation probe trial followed by removing the exemplar from sight and asking the child to find a successful tool. If the child used an item or combination of items successfully, it was pointed out immediately with descriptive feedback but not praise (e.g., "See, you hit that peg all the way down with the cotton spool.")

The training procedures' exemplars were not the regular tools as in baseline. Instead a different, and usually uncon-

ventional, exemplar tool was introduced in each session. The first appearance of a successful improvisation was reinforced with praise and descriptive feedback. Re-use of that item in the same session generated no comment. Re-use in subsequent sessions was noted only on the first occasion in a session, with descriptive feedback but not praise.

The combined use of train-to-generalize and train sufficient exemplars is illustrated in the following example. In training, the experimenter would demonstrate an alternative tool for a child, then would encourage the child to use it and would then socially reinforce its use. This sequence of events is an example of the train sufficient exemplars strategy. The exemplar tool was then removed from sight and the child was asked to come up with a new tool. If the child were successful, the experimenter would provide descriptive feedback and social reinforcement (e.g., "Good boy, you hit the peg all the way down with the piece of brick. That's great!"), which is an example of the train-to-generalize strategy.

Successful improvisations were defined by a performance criterion established for each task. The responses recorded were: picking up an item; attempting to use an item; and successful improvisation with an item.

The results of the study revealed that all but one child showed few improvisations without training. Comparisons bet-

ween baseline and training phases in each tool task trained showed greater increases in the diversity of improvisations after specific training was initiated. No generalization of training effects was noted in untrained tools. However, it must be acknowledged that the positive effects of training, where observed, were very small. Throughout 18 training sessions the best improvisation was eight containers by one child. The data also contained long periods in training where no new improvisations occurred. The Follow-Up data which was considered to show maintenance of improvisation performance by the authors revealed only six improvisations for five children in 11 cases where the tasks were trained, which were fewer improvisations than observed in baseline.

A number of difficulties can be noted in the improvisation training of Parsonson and Baer (1978). A novel for this experiment criterion was used to determine the reinforcement schedule (i.e., children were reinforced only the first time they used a particular object and never again). However, the next time that response occurred it was met with descriptive feedback for its first appearance. The very subtle difference between descriptive feedback alone and with praise may not have been discriminated by the children, so that the discrimination of the desired response may have been very difficult. This may have set conditions for perseveration on a successful item, which Parsonson and Baer indicated frequently occur-

red. In addition the children were always reinforced for completing the task with the exemplar both in baseline and in training phases. Furthermore they received reinforcement for attending the sessions (a small toy was given to each child after each session). With so many responses being reinforced the children may not have been able to discriminate between the desired response (i.e., novel tools) and the other responses (i.e., completing the task, completing the session early and getting a toy).

Using a novel for this experiment criterion for reinforcement may not have been adequate to bring the response under stimulus control, because the operant level may have been too low. Goetz and Baer (1973) for that reason had used a novel for this session criterion, reinforcing every new response in a session regardless of whether it had occurred before in other sessions. They found that the children still produced more new forms in later sessions than in earlier sessions. Of particular relevance to the improvisation problem is the maintenance of a successful response in a child's repertoire, in addition to the generation of new tools. In Parsonson and Baer's (1978) study, using a novel for this experiment criterion did not produce large gains in new improvisations. Perhaps the children had not been exposed to enough reinforcement to firmly establish novel responding in each session and to establish the discrimination between completing the task and/or session and searching out new alternatives. Parsonson and Baer (1978,

p. 379) suggested that a media trap may have occurred which entails the response of completing the task itself being more reinforcing to the children than praise from the experimenter. A media trap seems a feasible explanation for the low responding and perseveration, in view of the extremely sparse social reinforcement resulting from a novel for this experiment criterion. These problems suggest using train-to-generalize in a novel for this session criterion to capitalize on the higher operant level made available by the shift in criterion.

A second difficulty in the improvisation training was the return of each improvisation item back to the tray after successful use by the children. In so doing a problem situation could no longer be said to exist, because indeed there was a solution on the tray which the child had been reinforced for using (cf., Skinner, 1953). This may have been a strong test of generalized improvisation, but it was a discrimination that probably was difficult for the children to learn. That is, given they were reinforced for successfully using an item, that same response in the same situation was likely to recur. When it did, because the item was put back and was therefore still available, the child was given descriptive feedback (which was conceivably reinforcing to the child), so that same response was more likely to recur. The item was put back again and used again, this time the response received no attention, so that it was now on extinction, until its first use

in the next session, where it would be again given descriptive feedback. Such procedures constitute a complex schedule of intermittent reinforcement which may in and of itself produce perserveration. The resolution to this problem would simply be to keep a new problem situation at all times, by removing any items that were used successfully. While this is perhaps an easier discrimination and so a less rigid test of generalization, it does allow the children to be reinforced only for novel responses.

A third training difficulty of the Parsonson and Baer (1978) study lies in the way the train sufficient exemplars strategy was used. The strategy was to: "facilitate exploration of, and experimentation with, a diversity of apparently unlikely items in the probe item pool," (Parsonson & Baer, 1978, p. 378). Four points can be made about the procedures followed in training exemplars. One, the training exemplars came from a separate pool of items to which the child had no access (Parsonson & Baer, Note 2). This may have resulted in a lack of similarity between the exemplars and the potential tool items. Also the fact that the exemplars did not come from the tray may have led to an important salient common stimulus difference between training and generalization settings, i.e., the tray.

The second point relating to the manner of exemplar training is that presenting a complex exemplar intact, and reinforcing completion of the task with it, may have inadvertently

trained the child to take the exemplar from the experimenter and use it. This would not then have trained the child to search the tray for an item, pick one out, then use it. Also, by using complex exemplars the experimenters may have presented the children with an overly difficult discrimination and so limited the children's probable success.

A third point is that the essential characteristics of the exemplars were not actually trained. Parsonson and Baer (1978) indicated that the essential characteristics of the tools were learned by the children, only after many errors. The essential characteristics of the tools and exemplars could have been taught quickly by telling the children what to look for in each task (e.g., hardness, spaciousness, or suppleness) in a way that would have made the discrimination between potential tools and useless items more obvious. As a result of not training essential characteristics, initial attempts by the children often involved distractors that looked potentially effective. The children then proceeded in a trial-and-error way to discover the essential characteristics and eventually to discriminate between potentially effective and essentially useless alternative tools.

The fourth point is that the study continued to present exemplars throughout the training period, up to 37, which seems contrary to the "few", "sufficient" exemplars mentioned in Stokes and Baer (1977). Exemplars are to be presented until sufficient generalization occurs, then only probing should

continue. Either no generalization occurred, (but the authors claim it did, p. 375) or their criterion of sufficient generalization was a nebulous one. To test for generalization requires a phase where generalization programming has stopped as Stokes and Baer (1977, p. 350) indicated by definition: "Generalization will be considered to be the occurrence of relevant behavior under different, non training conditions (i.e., across subjects, settings, people, behaviors and/or time), without [italics mine] the scheduling of the same events in those conditions as had been scheduled in the training conditions."

These four points indicate four needs in the training of improvisation of tools; (a) the exemplars should come from the item pool, sharing the salient common stimulus of the tray; (b) reinforcement contingencies must be carefully placed on the desired behavior only, to establish clearly to the children what response is sought; (c) the essential characteristics of the tool items should be actively trained by enhancing the discrimination between potentially effective and essentially useless items; and (d) some criterion of "sufficient generalization" must be set to verify that generalization has occurred and to establish the number of exemplars that had to be trained to obtain generalization.

A major need which requires further examination is how discrimination and generalization are to be trained together. Further analysis in the following two sections will elucidate this issue.

Discrimination Training

As mentioned earlier (cf., p. 7), Skinner (1968) stated that discrimination is a central process in problem-solving. Parsonson and Baer (1978) have also indicated that in tool improvisation a discrimination must be established between potentially effective tools and useless items.

Discrimination problem solving has been extensively studied in a mediational model by Kendler (e.g., Kendler, 1971; Kendler, Glasman & Ward, 1972). Kendler et.al. (1972) studied discrimination shift behavior. For example, in the initial phase of a discrimination shift the subject may be confronted with a task involving two stimuli that vary on two dimensions, e.g., brightness and shape. In the first phase of discrimination black is correct and white is incorrect, irregardless of shape. After this discrimination is learned, a shift in the reinforcement contingencies occurs. In a reversal shift, the response contingencies are reversed, that is, white becomes correct and black incorrect, irregardless of shape. In an extradimensional shift, the reinforcement contingencies are shifted to the previously irrelevant dimension of shape. Whereas brightness initially was the critical dimension for reinforcement, now it becomes shape irregardless of brightness (Kendler, 1971).

Typically children below 12 years of age perform better on extradimensional shift problems than they do on reversal

shift problems. As a child matures, the ability to reverse a previously learned discrimination increases. The change in ease of reversal shift behaviors has been attributed to an ontogentic progression from a single-unit process in which the discrimination stimuli become directly connected to the choice response, to a complex mediational process in which the external cues are transformed into an implicit representational response ("black not white," to "white not black") which then cues the choice behavior (Kendler, 1971; Kendler et.al., 1972).

Kendler et.al., (1972) have developed a method of training preschool children in reversal-shift behavior. The strategy used essentially was to construct supplemental stimuli which then took advantage of the subject's stimulus control history. Verbal labeling and cue training were compared (Kendler et.al., 1972), as to their effectiveness at increasing subjects reversal shifts in sorting cards with curvilinear and rectilinear lines in a conceptually systematic way. The verbal labeling subjects were required to articulate distinctive verbal labels to the instances of each conceptual category with lines. For example, children were presented with a card which had a curvilinear line on it, the experimenter said "This is a picture of a curvey; What is its name?" When the subject responded correctly by saying "Curvey" the experimenter said "good!" The verbal label for rectilinear lines was "Pointy." The cue

training subjects were trained to match each line, before sorting it, with one of two critical cues, an acute angle or a curved line on a stimulus card, by tracing the lines with their fingers.

Both procedures were superior to a control condition, where no training occurred, in increasing reversal-shift performance. Verbal labeling was most effective, however. In a mediational model interpretation, the labels encouraged symbolic responses that represented the instances from each category and provided cues to guide the children's response (Kendler et.al., 1972), thereby improving performance. Alternatively, in a behavioral interpretation, the verbal labels facilitated the precurrent problem solving activity of attention, by serving as supplemental stimuli and thereby taking advantage of previously established stimulus control. Generalization to other problems occurred because the verbal labels were salient common stimuli, as described earlier in the mediate generalization strategy.

Generalized Improvisation and Discrimination Behavior

The improvisation of tools problem requires that children select a functionally successful tool out of a pool of effective and useless items. The children must discriminate between useless items and potentially useful tools. Conceptually this problem seems similar to a discrimination-shift problem. The children, given familiarity with the task, have been reinforced in the past for completing the task with a specific tool (e.g.,

a hammer). Now they must complete the task using an item that was irrelevant to the task before (e.g., an ash tray). They must select their new tool on a functional dimension not a shape dimension which would constitute an extradimensional shift. Parsonson and Baer (1978, p. 378) noted that the children initially chose items that looked effective, that is, they considered shape as an essential dimension for reinforcement.

Given the conceptual similarities between the improvisation problem and discrimination-shift problem and the success of the Kendler et.al. (1972) training, using verbal labeling seemed like an expeditious route to increase generalized improvisation. It was particularly attractive as well in that it fit into the generalization strategy called mediate generalization (Stokes and Baer, 1977). Furthermore, using verbal behavior to facilitate generalization corresponded to Skinner's (1968) recommendation that constructing supplemental stimuli through instructions gains attention and facilitates problem solving.

Hypotheses

Systematic analysis of tool improvisation in children was undertaken, since it was viewed as a valuable skill, because of the natural reinforcement given to "clever" children in their play interactions with the items in their environment. Furthermore tool improvisation was viewed as a basic skill for children to possess in future problem solving situations in

which generalization was important. Therefore the establishment of the most effective generalization technique might have broad implications for training problem solving.

The present research trained preschool children in the generalized improvisation of tools (hammers and containers), by employing a precurrent problem solving skill: attention to essential characteristics. Attention to essential characteristics was defined as discriminating between potentially effective and essentially useless items. The following strategies were examined for efficiency of training and amount of generalized improvisation. Strategy one, train-to-generalize, provided descriptive social reinforcement for novel responses. Reinforcement was provided on a novel for this session basis. However, in all four training strategies, any novel improvisation by a child in a given session was removed to maintain the essential problem characteristics, thereby functionally constituting a novel for this experiment criterion. Strategy two, train sufficient exemplars, provided exemplars of searching out and using alternative tools from a tray of items with no descriptive social reinforcement. Strategy three, sufficient exemplars plus descriptive praise, provided a combination of the conditions of strategies one and two. Strategy four, mediate generalization, provided verbal labels of the functional discrimination between effective tools and useless items in the presentation of exemplars, plus descriptive social reinforcement.

It was predicted that the first strategy, train-to-generalize, alone would result in the least amounts of generalization because a low operant level of the behavior desired might not allow any stimulus control to be established. Strategy number two, train sufficient exemplars, alone would produce slightly more generalization because of the use of diverse exemplars which would model for the children the searching out and use of items from the tray. However, it would also be limited because no differential reinforcement was used to maintain responding. Sufficient exemplars plus descriptive praise, should produce higher, moderate levels of novel improvisation, as it did in Parsonson and Baer's(1978) study, because the exemplar training would be used with differential reinforcement. The exemplars should increase the likelihood of occurrence of target behaviors which once they occurred, would be strengthened by the descriptive social reinforcement. The most successful strategy was predicted to be mediate generalization, because the verbal behavior should provide cues to facilitate discrimination of potential tools from essentially useless items. As well, the verbal behavior should produce salient common stimuli which would facilitate generalization across sessions and possibly across tools. In addition, this strategy also utilized exemplars to elicit the target behaviors, and descriptive social reinforcement to increase the likelihood of the target behavior's recurrence.

Generalization across tasks if it occurred, was expected to be less than generalization within a task trained, because of the strong topographical differences in the two tasks. Also Parsonson and Baer (1978) had demonstrated no across task generalization.

The behavioral measures used to assess the predictions stated above were; (a) the cumulative number of improvisations per session; (b) the percentage of successful discriminations per session; (c) the mean rate of improvisation per session; and (d) the number of exemplars required to achieve generalization (used for all but train-to-generalize where exemplars were absent). The rationale for the first measure was that the cumulative number of improvisations made by the child in each session would demonstrate whether the child had or had not learned to make new alternative tools. Also cumulative number of improvisations was used by Parsonson and Baer (1978) and so would facilitate comparisons of results. The rationale for the second measure was that generalization of this nature could only be properly assessed, when the number of useless items used in relation to the number of improvisations was known, because new alternatives may only be accidental successes, the result of trial-and-error interactions with all the available items. Therefore the production of many novel alternatives must be accompanied by a decreased use of useless items which would allow the establishment of sufficient generalization

(cf., Research Design and Data Analysis section). The rationale for the third measure was that rate of improvisation per session was important to compare the efficiency of the four training strategies because of the varying number of improvisations and sessions in training that were likely. The rationale for the final measure was that sufficient exemplars training by definition meant training using only a few response examples, not all of the example tools. Therefore the number of exemplars had to be known to establish what was a sufficient number of exemplars in these problem situations.

Method

Subjects and Setting. The subjects were 10 preschool children, ages 31 mos. to 55 mos.; the mean age was 50 mos., the S.D. was 7.6 mos. The children attended the Greta Brown Unit Day Care Centre in the low income, core area of Winnipeg. The children were selected by the staff at the day care centre on the criterion that they were above age 24 mos. and below age 60 mos. Letters explaining the research to be conducted were sent to 15 parents, or guardians, of which 10 agreed to allow their child to participate in the project.

Before selecting preschool children for this research, kindergarten children had been recruited and the training initiated with them. Kindergarten children had been selected because the older children (one 5 yr. old and one 6 yr. old)

in the Parsonson and Baer (1978) study showed the greatest and most consistent responses to training. However, with the considerable changes made to the problem situation here, the kindergarten children displayed a maximum rate of improvisation with 100 percent discrimination, without training. Younger children were then sought to ensure a lack of improvisation behavior in the tasks selected for the research.

Once the children had been selected and their parents' permission to participate received, they were randomly assigned to one of the four training strategies. Individual 10 minute experimental sessions were conducted once per day, 5 days per week, in a room 2.13 m. by 1.83 m., which was adjacent to the day playroom. The experimenter sat opposite from the child at a table 91 cm by 49 cm, with the tools on the table between them. The table was used (unlike Parsonson and Baer who worked on the floor with the children) for two reasons. First, the day care centre director urged that a table be used because the children were accustomed to working with adults in this way and would probably thus cooperate better. Second, the room did not have a carpeted floor which therefore made it unreasonable to have the children and particularly the experimenter on the floor. The marbles were kept on a dinner tray on the table, so there were few times when marbles rolling onto the floor constituted a problem.

The session lengths were the same as in Parsonson and Baer

(1978), that is, 5 minutes for each task. It was uncertain at the outset of this study if the time alotted would be long enough, or if the children would be seriously hampered by a lack of proficient motor skills. If it was deemed necessary the experimenter was ready to increase the session length for whichever task needed it. Before session 23 it appeared that most children had perfect discrimination, used up all 5 minutes, but did not use all the potential tools in the container task. The session length for containers was then increased by 5 minutes. No change was made to the length of the hammer session.

Equipment and Performance Criteria. Two tasks (hammers and containers) were chosen from Parsonson and Baer's(1978) study, as well as some task items and certain of the success criteria. Numerous potential tool items and distractors were added to those of Parsonson and Baer, because many of their potential tools were combinations of two other potential tools. In many cases these combinations considered as new tools were trivially different from the original or each other and did not appear functionally advantageous to task performance. As well, if the potential tools were simple, it was thought that then the children were more likely to be successful and so come under the reinforcement contingencies sooner. Much of the perseveration noted by Parsonson and Baer could have been due to the experimenters' view that more combinations of items were possible, while the child's view was that

the task was to be completed (because doing so had been reinforced in the past) and the problem situation was the same as before, because a previous solution was available. The present research attempted to keep a "true" problem situation by always removing successful improvisations. Furthermore simple solutions would increase the likelihood of success and thereby reinforcement. (See Appendix A for a complete list of items used).

The hammer and container tasks were chosen and the shoelace task omitted because of the lack of suitable improvisation items and distractors for the shoelace task. In addition, the children may not have been familiar with the shoelacing response, as noted by Parsonson and Baer. Lack of facility with shoelacing hampers one of the initial conditions in setting up a problem situation, i.e., that there be a response in repertoire which is being prevented from being emitted (Skinner, 1968; Bijou, 1976).

The items, both potential tools and useless items, were placed on a dinner tray, on the table in front of the child, between child and experimenter. Though Parsonson and Baer used five potential tools and 10 distractors, each tray here contained five potential tools and five useless items. Examples of these materials are provided in Table 1. The number of items was selected arbitrarily to fit the limited size of the tray and table surfaces. As well, 10 items were thought to be a manageable number for the children. An equal number

Table 1

Examples of Materials Used: Exemplars; Potential Tools; and Distractors in Two Tool Tasks

Tool Use Category	Tool Task	
	Hammers	Containers
	Playskool Cobbler Bench	80 marbles
Baseline Exemplar	wooden hammer	plastic bag
Training Exemplar*	door knob assembly	egg carton
	plasticene ball	child's shoe
	ashtray	baseball glove
	adjustable wrench	pastry bag
	"D" size battery	man's glove
Potential Tool*	garden shovel	man's mitt
	croquette ball	hand towel
	salt shaker	ice cube tray
	paint brush	frisbee
	padlock	nite cap
Distractors	soft rubber hammer	ice cream cone
	6 inch plastic ruler	mesh purse
	plastic wrench	35 mm film can.
	pencil	cup-cake cup
	"AA" size battery	jar lid

*exact exemplars and potential tools used for each child depended on each child's improvisation success, i.e., some children may have used some of these items as improvisations, therefore, other items were chosen as exemplars. Similarly potential tools varied depending upon exemplar selection and improvisation success.

of potential tools and distractors was chosen in order to equalize the probability of choosing either type of item by chance selection. This condition existed only before exemplars were trained, however.

Improvisation Tasks. The two tasks were, pounding on a poundingtoy and storing 80 marbles. Improvisation was of hammers and containers. The materials for each task presentation included a tray, five potentially successful tools and five useless (distractor) items.

Hammers. The hammer task was to pound down a wooden peg in a Cobbler's Bench toy. An effective hammer was defined as any item that would pound one peg to within 3.2 mm of the bench top.

Containers. The containers task required putting 80 marbles in an item that would hold them without loss, when picked up off the table.

Behavioral Definitions

Parsonson and Baer (1978) recorded the number of item pick-ups, the number of task completion attempts, and the number of new improvisations, per session. The three response classes were accorded different values with pick-ups receiving the lowest and improvisations the highest value. The final data of concern was the highest level of interaction with an item,

which would appear to indicate that little of the other two categories was recorded or that they were not useful, since the data were not presented in the published report. On this basis and for the reasons previously mentioned in the Hypotheses section, only improvisations, distractor uses and exemplars were defined and recorded in the present study. From these data, the four response measures: (a) cumulative improvisations; (b) percent successful discrimination; (c) rate of improvisation; and (d) the number of exemplars trained; could be calculated to assess the amount of genuine tool improvisation.

Improvisation (I); Recorded when an item was used to meet the performance criterion.

Distractor (D); Recorded when an attempt to complete the task was made with an item that did not meet the performance criterion. This scoring included both what had been defined as distractors or potential tools, if the performance criterion had not been reached.

Exemplar (E); Recorded at the beginning of each session in which the experimenter used any item to demonstrate an example of a successful tool or distractor.

Data Collection Procedure

Recording interactions with items involved writing down when an item had been interacted with in the ways listed above.

One complete interaction was defined as, making physical contact with an item (both potential tools and useless items) by lifting it off the tray and making an attempt to complete the task with a distractor, or successfully completing the task with a potential tool, and putting the item down again. The experimenter was the primary observer.

Inter-Observer Reliability. Checks on reliability of observation were made by another observer in addition to the primary observer, once per week throughout the baseline and training sessions. The reliability observer was aware of the general purpose of the study, and the experimenter was aware when the reliability checks were being made. Inter-observer reliability scores were obtained by comparing the data sheets of the two observers and scoring an agreement or a disagreement for each tool interaction recorded as an improvisation or distractor. The total agreements were divided by the total agreements plus disagreements and multiplied by 100. The range of reliability scores obtained for the container task was 0% to 100% with a mean reliability score of 98%. The zero reliability score occurred in the first week of the study when a child used a small plastic bag to store marbles. All of the marbles were in the bag; however, when she attempted to move the bag the marbles fell out. As a result the experimenter scored a distractor attempt and the reliability observer scored an improvisation. This was the only item interaction

in that session resulting in a 0% agreement score. The 0% score was highly unrepresentative; but since the reliability coefficient for train sufficient exemplars was still very high, 91% (see below), it was included without further worry. The range of reliability scores for the hammer task was 83% to 100%, with a mean score of 99%.

Reliability scores were calculated for each training strategy. The range of reliability scores for the train-to-generalize strategy in containers and hammers was zero, with a mean score of 100% for each task. The range of reliability scores for the train sufficient exemplars strategy in hammers was zero with a mean score of 100%. The range of scores in containers was 0% to 100%, with a mean score of 91%. The range of reliability scores for the sufficient exemplars and descriptive praise strategy for hammers was 83% to 100%, with a mean score of 98%. The range scores in containers was zero, with a mean score of 100%. In the mediate generalization strategy the range of reliability scores for both hammers and containers was zero, with a mean score of 100% for each task.

Training Procedure

Session Introduction. Each child was brought to the training room by the experimenter and was seated opposite him at the small table. The order in which the children were brought to the training room was not systematically determined. Because the children played outside with the other 50 plus chil-

THE UNIVERSITY OF MANITOBA LIBRARIES

dren of the day care centre, it was more efficient to take children as they were available. Also the experimenter did not want to remove the child from a reinforcing play setting because the resulting emotion due to the time-out, and competing reinforcement would probably have affected the training session performance.

The tasks were presented to each child in a consistent order and no counterbalancing of the task trained first to assess task order effects within training strategies was done. No counterbalancing was used because; (a) the two tasks could be considered independent; as Parsonson and Baer demonstrated no generalization across tasks; (b) the children, in addition to being in a multiple baseline across tasks design, were yoked with at least one other child in a multiple baseline across subjects design, therefore it was necessary for each group of children to be experiencing conditions in the same order (a basic assumption of a multiple baseline across subjects design); (c) counterbalancing in Parsonson and Baer showed no differential effects; and (d) the already complex design had the experimenter completely involved in arranging conditions (See Table 2 for a summary of the research design).

Baseline. Once the child was seated in the training room the first of the day's two tool tasks was introduced by picking the exemplar (the regular tool) from the tray of items and then describing and demonstrating criterion performance,

Table 2

Experimental Conditions For Each Subject

Strategy Used	Subject	Order of Training	Sessions per Phase: Baseline	Training	Remedial Training*	Follow-Up
Train-to-Generalize	Jason	Hammers	1-5	6-25	-	-
		Containers	1-11	12-17	18-29	30-31
	Michelle	Hammers	1-14	15-26	-	-
		Containers	1-16	17-20	21-27	28-30
Train Sufficient Exemplars	Sylvia	Hammers	1-4	5-18	-	19-23
		Containers	1-10	11-24	-	25-27
	Aysa	Hammers	1-9	10-15	-	16-22
		Containers	1-13	14-27	-	28-30
Sufficient Exemplars + Descriptive Praise	Wendy	Containers	1-5	6-25	-	26-28
		Hammers	1-18	19-26	-	27-29
	Che	Containers	1-9	10-29	-	30
		Hammers	1-13	14-19	-	20-22
	Daniela	Containers	1-14	15-27	-	28-31
		Hammers	1-25	-	-	-
Mediate Generalization	Delia	Hammers	1-4	5-19	-	20-22
		Containers	1-8	9-29	-	30
	Alvin	Hammers	1-8	9-16	-	17-22
		Containers	1-12	13-24	-	25-28
	Tanya	Hammers	1-14	15-21	-	-
		Containers	1-19	20-21	-	-

*Mediate Generalization

(e.g., "This is a hammer, and it can pound pegs right down.") The baseline exemplars were a wooden toy hammer and a 26.7 cm by 20.3 cm plastic bag with a draw string. After the demonstration of criterion performance the child was then given the exemplar tool and was invited to complete the task. Verbal encouragement, was given to complete the task if necessary, but there were no further comments or social reinforcement.

An improvisation probe trial followed, by withdrawing the exemplar from sight and asking the child to complete the task using a new tool, (e.g., "Now, look at the things on the tray and find something that you can hit a peg right down with.") The child was given 1 minute to make a response before the trial was terminated and the next tool introduced.

Successful completion of the task was met by removing the item used to improvise so as to maintain the essential characteristic of the problem situation. The child was then asked to try again, (e.g., "Now, can you find something else that you can hit a peg right down with?") The trials continued until the child announced completion (e.g., "I can't find any more hammers.") or 1 minute had elapsed since the last attempt with any item had been made, or 5 minutes had elapsed since the child was first shown a tool exemplar.

Any improvisations by a child were removed immediately in the session in which they were produced and were not returned to the tray in the next session(s). This was done

in order to maintain the essential problem setting, where no previous solution was present. A new tool item was added in the next session to keep the number of potential tools and distractors constant at the beginning of each session. Permanent removal of a child's improvisations thus clearly established a novel for this session criterion, without the child having to learn such by trial-and-error. Because the improvisations were never returned to the tray, this forced a novel for this experiment criterion to be established, but without trial-and-error learning of the contingencies. In each session as the child improvised and those items were removed, the ratio of potential tool items to distractors decreased, essentially making the task more difficult because there were more distractors to choose from in the discrimination process. Continued improvisation success also continued to increase the difficulty of discriminating effective tool items from useless items in terms of chance responding. A new item was not added until the next session because the addition of an item during the session would have singled it out for use, thereby making it an experimenter selected exemplar rather than a child selected solution.

Improvisation Training

Train-to-Generalize. As in baseline the children were shown the baseline exemplar and were asked to come up with an alternative tool. However, now successful completion of the

task was immediately met with descriptive social reinforcement, (e.g., "See you pounded the peg right down using the brick, that's great!"). The improvisation item was then removed, so that the child could not use it again and so face extinction, because no experimenter reinforcement would be forthcoming. This would then clearly establish what response was desired, that is, completion of the task with a new tool, both in that session and in subsequent sessions.

Train Sufficient Exemplars. After baseline sessions a new exemplar tool was used in each session. The experimenter took an item from the tray, performed the task, replaced the item and then the child was encouraged to do the same with the item. Once the child had completed the task, the exemplar item was removed and the child was asked to find an alternate tool.

The selection and presentation of exemplars attempted to cover the entire array of tools available (See appendix B for a list of exemplars used for each child). That is, the items were considered as to how they were used; either by a handle or not for hammers or either being a flat container or an open container. The exemplars changed between each type of tool usage alternately from session to session.

Sufficient Exemplars Plus Descriptive Praise. Exemplars were presented as in the previous condition. Now, however, successful completion of the task with an improvisation item was reinforced using descriptive social reinforcement. This

was the only change from the sufficient exemplar alone procedures.

Mediate Generalization. Exemplars and descriptive social reinforcement were presented as in the previous condition. In addition verbal labels were used to facilitate the discrimination of the tool's essential characteristics and to facilitate generalization. Once an exemplar had been presented the experimenter said, "This is like a hammer, because it is hard just like a hammer." Then picking up a distractor said, "This isn't like a hammer because it's too soft." The experimenter then had the child use the exemplar and asked why the item had worked (e.g., Experimenter; "Why does the rock work?" Child; "Because it's hard!" Experimenter: "Very good, that's right!"). The exemplar of a potential tool and a distractor were then both removed and the child was asked to find an alternate tool. The ratio of potential tools to distractors remained equal (now 4:4), until the child successfully improvised in which case improvisations were removed as in the previous strategies.

The slight variation in number of starting items in the different strategies (10 for train-to-generalize, 9 for sufficient exemplars and sufficient exemplars plus praise, and 8 for mediate generalization), as well as the varying ratios of potential tools to useless items represents a minor confounding in the procedures. Presumably when the ratio of

potential tools to useless items is equal, by chance there is equal probability of either a correct or an incorrect response. The discrimination becomes slightly more difficult as the ratio of potential tools to useless items gets smaller. Thus, from this perspective successful generalization when obtained may represent a conservative estimate. However, the number of total items decreases which possibly produces an easier problem situation because there are less items from which to choose. As mentioned previously (cf., p. 36) Parsonson and Baer (1978) had five potential tools and 10 distractor items at all times. It would seem then that the best combination of items and the best number to choose from to best test generalization is unclear. While one situation may be an easier overall situation, it also constitutes a more difficult discrimination, according to chance responding. Since chance responding is not what is to be acquired, this potential confounding was not viewed as serious. Nevertheless, research is necessary to determine its impact.

Reinforcement Only. Once a child reached the sufficient generalization criterion, the sessions which followed involved asking the child to find an alternate tool without any prior demonstration of criterion performance with an exemplar. Correct improvisations were met with descriptive feedback and social reinforcement, and then the item was removed. This final phase was intended to assess whether generalized improvisation had

been established for the exemplar training strategies. The reinforcement contingency was also intended to replicate what a child might expect in the natural setting. The skill of tool improvisation would thereby be maintained in the children, something considered as a valuable skill. This final condition was similar to the Follow-Up phase used in Parsonson and Baer (1978), where the children received descriptive feedback but no praise for successful improvisations. The conditions both here and in the Follow-Up of Parsonson and Baer followed training immediately and assessed only the short-term maintenance of improvisation.

The reinforcement only conditions were not applied in the cases where train-to-generalize had successfully trained sufficient generalization, because there was little substantial difference in the contingencies, i.e., the only extra condition in train-to-generalize was the demonstration of the regular tool. Since the exemplar was not considered to be the potent feature of the train-to-generalize strategy (cf., p. 16) it was not deemed necessary to test for genuine generalization as per Stokes and Baer's definition (cf., p. 25) by probing without the exemplar as in the cases of exemplar training. The number of sessions in this phase depended largely on how many tools were still available, although if possible three sessions were used to clearly establish if the trend was being maintained (Hersen & Barlow, 1976).

Multiple Baseline Reversal Design

The research design was a multiple baseline across tasks

and across children, with several different reversal conditions which varied in each strategy. The multiple baseline technique was used to measure baseline improvisations on both tasks and on all children. Within each strategy the intervention occurred in one task for the first child while monitoring the baseline conditions for all the other tasks and children. Any changes in the level of improvisation were noted and evaluated against the baseline levels of improvisation in other children or on the other task. Then, the same experimental condition was applied to the second task and/or the second child and so on until all tasks and all children had had the training intervention applied (Baer, Wolf & Risely, 1968).

Once sufficient generalization had been reached a third condition the reinforcement only, (also referred to as follow-up or generalization probing stage), was introduced subsequently to all tasks and all children. This condition allowed the evaluation of sufficient generalized improvisation, where exemplars were taught. As mentioned previously (cf., p. 50) this condition, which was different from baseline in that it involved descriptive feedback and social reinforcement, was used because maintenance of the improvisation skill was desired (Hersen & Barlow, p. 225). In those cases where little or no improvisation had been developed, the third condition was not generalization probing but was mediate generalization. This strategy was conceptually viewed as the superior strategy as discussed earlier. In addition, very early in its implementation it produced the most consistent

effect across the two tasks being trained. The consistency of effect made mediate generalization an even more attractive strategy to use if any of the other strategies failed. Moreover, as mediate generalization involved more training conditions than the other strategies, basically being an increment on the other strategies, it was the logical replacement to develop improvisation, when one of the less complex strategies had failed. Once a sufficient amount of generalized improvisation had been reached, the final condition, the reinforcement only phase, was implemented.

Results

Dependent Measures

In all training strategies the cumulative number of improvisations was graphed across sessions. As well the percent successful discrimination was calculated by totalling the number of improvisations for each session then dividing by the total number of improvisations plus distractor uses and multiplying by 100. The discrimination measure was also graphed across sessions. It thereby yielded a continuous percentage measure of the ratio of successful improvisations to distractor uses plus improvisations, that is total attempts with any item. This index of successful discrimination in comparison with cumulative number of improvisations would indicate if genuine generalized improvisations were occurring, or if trial-and-error

responding were producing the cumulative improvisations. These two measures were always considered when a change from baseline conditions to training conditions was made for any subject.

The mean rate of improvisation per session was calculated by dividing the total improvisation of a subject in a given phase by the total number of actual training sessions in that phase. This measure yielded a figure which could be used to compare effectiveness across training strategies. Finally, the number of exemplars required to achieve generalized improvisation was used to compare the exemplar strategies and establish "sufficient" exemplars.

Sufficient Generalization Criterion. A criterion of sufficient generalization was established using the first two measures above. Once a child had used all of the potential tools available in a session with a percent discrimination level of 80% or better for three consecutive sessions in the hammer task, generalization was considered "sufficient". For containers the criterion used was, use of all potential tools available in a session with a percent discrimination level of 80% or better in two consecutive sessions. The criteria were established according to the following considerations; (a) no precedent research had clearly indicated any levels of generalization considered sufficient (Baer et. al. (1967) and Guess et.al. (1968) appeared to use 80% correct imitation, therefore an 80% discrimination level was adopted here); (b) all

or nearly all the potential tool items had to have been used in order to make the percent discrimination level meaningful, because it was possible to achieve a 100% level and only be producing one improvisation, whereas only one incorrect attempt followed by five correct responses would reduce the discrimination level by 20%, so that 80% success is in fact a stringent criterion; and (c) three sessions were adopted to measure criterion performance based on Herson and Barlow's (1976) recommendations that three successively increasing or decreasing data points were required to establish a trend in the data. It subsequently became clear that due to the session length being too short for containers, longer training phases would be required to reach criterion. In order to approximately equalize the length of task training conditions, because of the potential shortage of viable tool items and distractors, and due to the possibility of children leaving the project with longer training periods, a shorter two day criterion was used for containers.

Distractors were only scored and counted if they occurred before all the potential tools had been used; after that the child was probably only making attempts because of being prompted to look for more tools. The 80% as opposed to a more stringent 100%, criterion seems appropriate because in addition to the research precedents of Baer et.al. (1967) and Guess et.al. (1968), many of the distractors could really only be established as such by actually testing them, and as mentioned previously, only one incorrect response reduced discrimination to 80%.

The improvisation data for both tasks have been presented in graphs displaying the cumulative number of improvisations per session for each child. The data are also presented in graphs displaying the percent successful discrimination per session for each child. These graphs are first shown to establish the multiple baseline across tasks design. The same data are also presented later in the multiple baseline across subjects design for the four training strategies.

The maximum number of improvisations for each child in each tool class was dependent on how soon the child reached sufficient generalization and the reinforcement only phase. The maximum number of items collected that could be used as exemplars or tool improvisations was 85 for containers and 53 for hammers; the ordinates on the graphs relfect the relative maximums reached by each child. The rate at which the children reached sufficient generalization was seen as a more important measure of strategy success and not overall maximum number of improvisations, because it was possible to improvise poorly and so remain in training longer thus increasing the eventual maximum number of improvisations. Some children were able to exceed these theoretical maximums by forcing items defined as distractors, by the experimenter's ability to use them as improvisations, to complete the tasks. This usually required an excessive amount of force, for hammers, or extreme agility at stacking marbles (or at least more strength or

agility than employed by the experimenter). These uses were scored as improvisations for that child, but were not re-defined for the other children, because it was thought they would be unable to use the extremes necessary. Indeed this was born out by some other children's unsucessful attempts with the same item(s).

Multiple Baseline Across Tasks Data

Train-to-Generalize.

Jason. Figure 1 presents data on Jason's improvisation within and across tool tasks. The cumulative number of improvisations graph shows that during baseline Jason produced three improvisations in the hammer task. Without training improvisations became zero. Baseline levels of discrimination were generally below 50% with only one or two improvisations, indicating trial-and-error success. Hammers began receiving training first where a slight increase in improvisations was noted, then no improvisations were produced for four sessions. Next a steady increase of two improvisations per day began. Initially in training the discrimination was 50%, rising to 100% and falling to zero. The discrimination jumped sharply in session 15 when improvisation began at the steady two per day rate. Discrimination began to fall again at session 17. However by session 19 it had begun to rise, reaching a high level of between 80% and 100%, by session 22 and attaining the criterion of sufficient generalization by session 25.

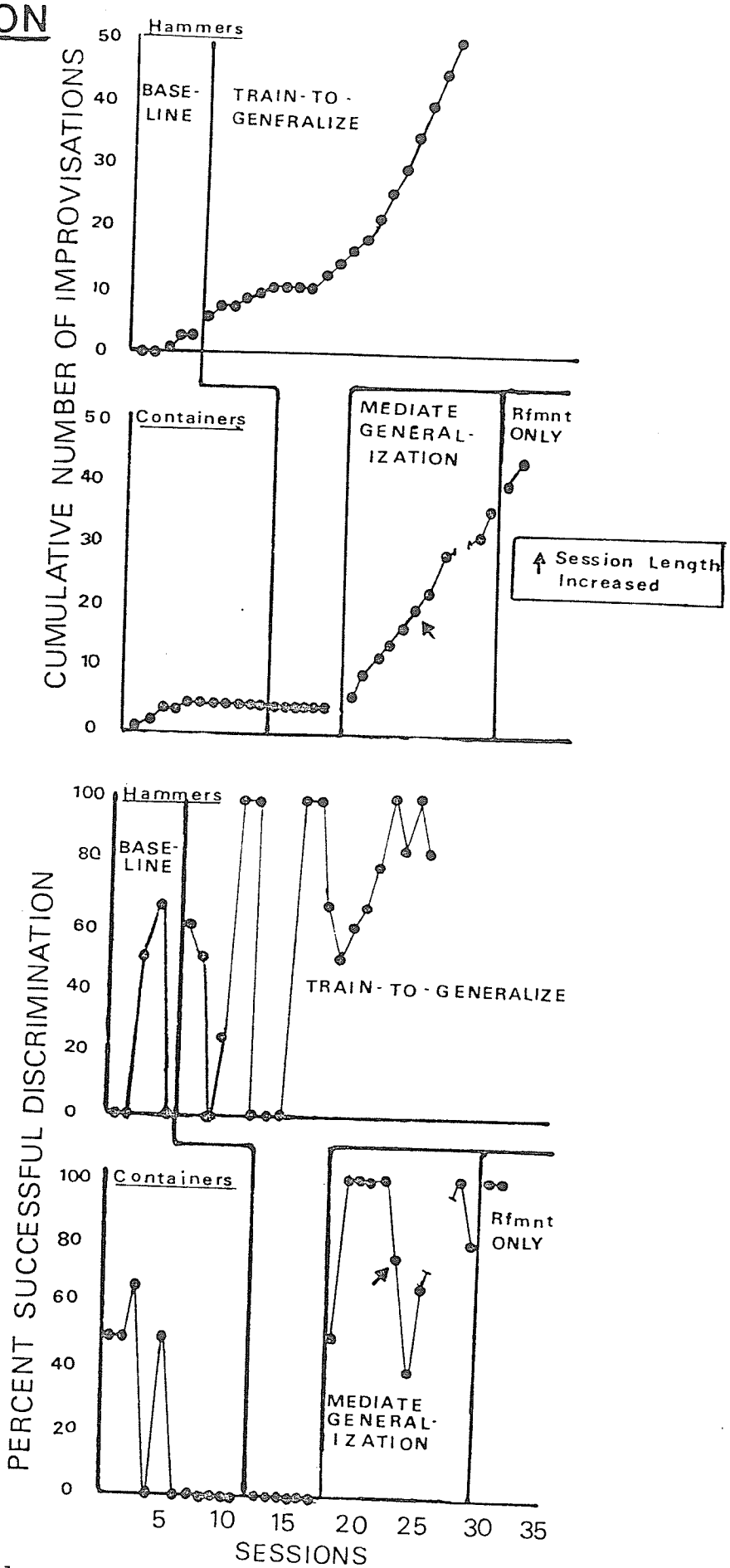


Figure 1. Cumulative number of improvisations and percent successful discrimination for Jason in hammers and containers.

Containers received training six sessions after hammers. There had been no generalization of the train-to-generalize effects in containers although there was no significant control in hammers and little was expected. Improvisation training began with improvisation at zero in baseline. Training had no effect on either improvisation or discrimination, both remained at zero. Because the operant level was zero, the differential reinforcement of train-to-generalize could not affect the behavior. Mediate generalization training was begun for containers in session 18; discrimination rose to 100% in two sessions and daily improvisations rose steadily. In session 23 when the session length was increased, the rate of improvisation quickly reached sufficient generalization and was maintained for 2 sessions in the reinforcement only phase before Jason left on vacation. Discrimination, however, dropped off sharply for three sessions after the length was increased. Increased time may have fostered increased responding to terminate the situation. Percent discrimination did rise to a high level again and remained at 100% in the reinforcement only phase. The definite increase in slope for cumulative number of hammer improvisations (rate changed abruptly from two per session to five per session) and discrimination one session after mediate generalization training had begun in containers strongly suggests that the training effects had generalized across tasks.

Michelle. Figure 2 presents the improvisation data for Michelle. Baseline sessions indicated steady increases in hammer improvisation, though the rate was only half the maximum rate. Discrimination was generally high but quite variable, from 0% to 100%. Training was initiated, after four weak sessions (where Michelle only produced three improvisations) in order to attempt to raise her daily rate of improvisation. Train-to-generalize essentially had no effect at increasing her daily rate of improvisation; her rate stabilized at one per day for 9 days. Discrimination was quite consistently 100%, except for two low days. In session 23 hammer improvisations reached a maximum daily rate with consistent 100% discrimination, but only after mediate generalization training had begun in containers.

Baseline sessions in the container task revealed few consistent improvisations except for four days starting from session seven. The rate of improvisation was zero for session 13 to 16 when training began. Throughout baseline there was extreme variability in Michelle's container discriminations, although discrimination became zero by session 13. Train-to-generalize was started in session 16, with no initial effect in either number of improvisations or successful discrimination. Two more train-to-generalize sessions were carried out, in which two improvisations were made with 100% discrimination. At this point given the poor performance of

MICHELLE

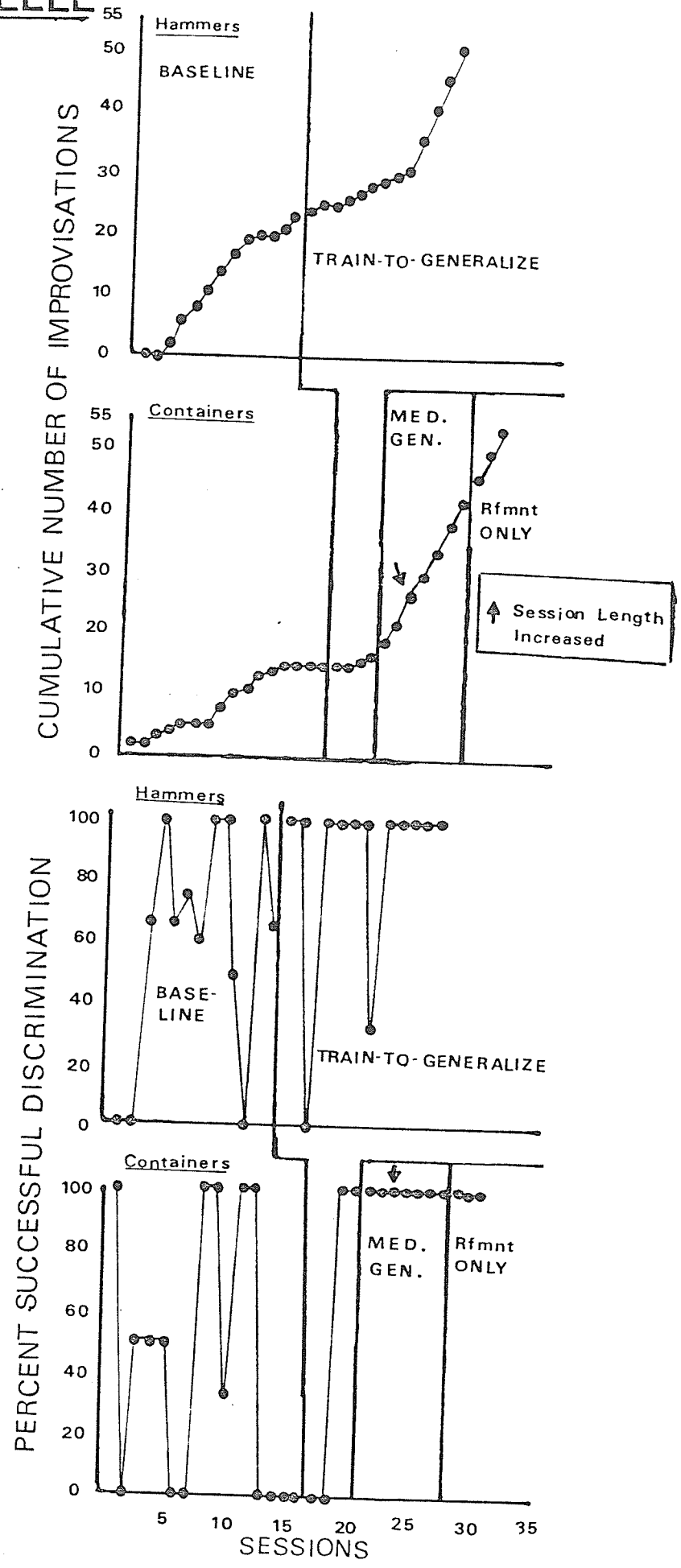


Figure 2. Cumulative number of improvisations and percent successful discrimination for Michelle in hammers and containers.

train-to-generalize in Jason's containers data, and in hammers for Michelle, mediate generalization training was started in session 21 to replicate the substantial effects shown in Jason's data when mediate generalization was used and to increase Michelle's improvisations. Mediate generalization had a dramatic effect on number of improvisations with an even larger daily increase noted after the session length was increased. Discrimination remained remarkably stable at a perfect 100%. High levels of improvisation and discrimination remained through the reinforcement only sessions. As briefly alluded to before two days after mediate generalization training was started in containers a sharp rise was noted in hammer improvisations, with corresponding 100% discrimination, reaching sufficient generalization in four days, strongly suggesting generalization of training effects across tasks.

Train Sufficient Exemplars

Sylvia. Figure 3 presents the cumulative number of improvisations and percent successful discrimination data for Sylvia. Baseline improvisations of hammers was zero with a zero discrimination level. Train sufficient exemplars began for hammers in session five. The effect of training was only seen two days after training began. The daily rate of improvisation showed a steady, but low increase up to session 15, where it rose to a maximum rate. Discrimination data showed a variable increase until session 15 where it remained at a level at or above 80%. When sufficient generalization was reached, the

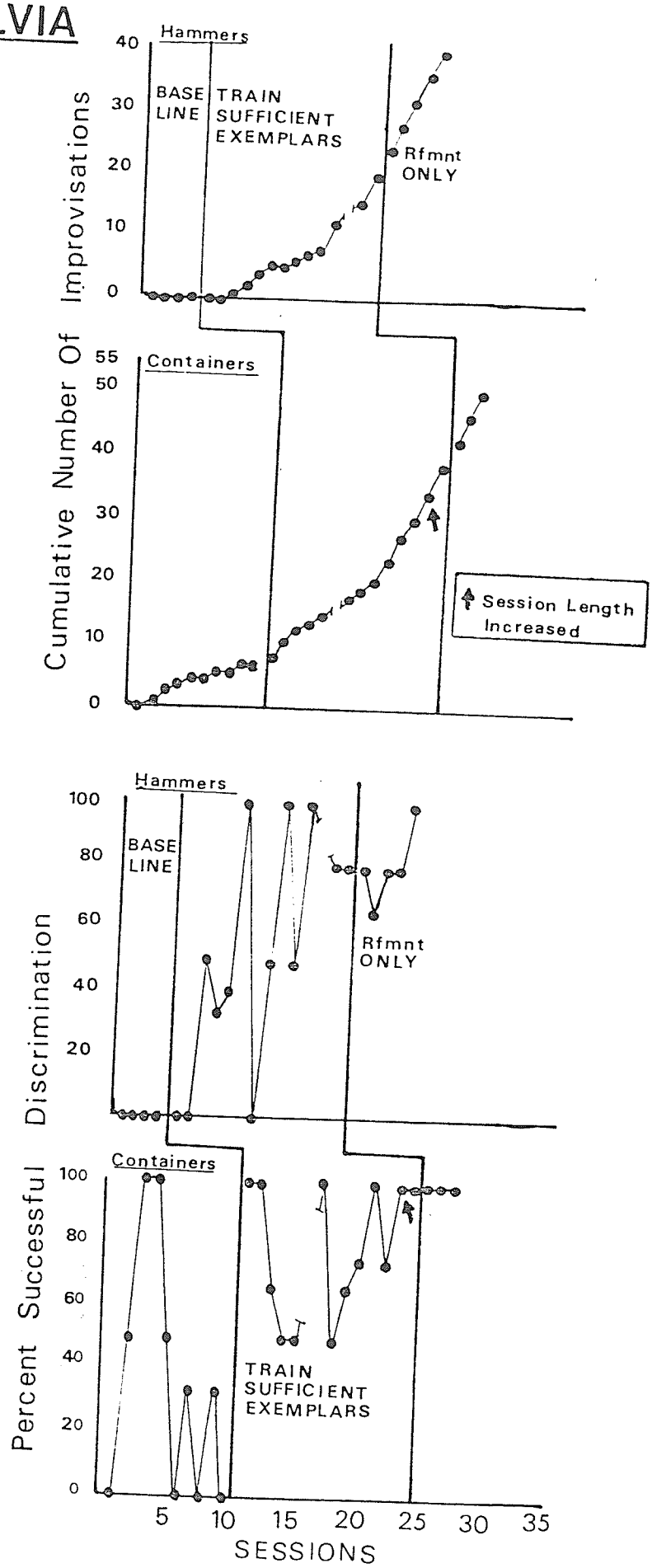


Figure 3. Cumulative number of improvisations and percent successful discrimination for Sylvia in hammers and containers.

rate of improvisation remained high in the reinforcement only phase. Though discrimination dropped off to 66% for one session, it generally remained at or above 80%.

Sylvia's container data show few improvisations in baseline. Training in containers began before strong control of training was demonstrated in hammers, limiting the interpretation of across task generalization. It was not known, at that time, if train sufficient exemplars would ever show strong control in either task; therefore, the limitation of subject time became a significant concern in deciding to implement training, because remedial training could have been required. Container discrimination in baseline rose to 100%, then fell sharply to a level between zero and 33%. This low discrimination level was accompanied by low (zero or one) daily improvisation. Training then began in session 11, where improvisation rate showed a significant increase. Discrimination rose sharply, remaining mostly above 50% until it stabilized at 100% for the last two sessions with maximum improvisation. Discrimination and improvisation remained high and stable through three reinforcement only sessions.

Aysa. Figure 4 presents the improvisation data for Aysa. Baseline improvisations of hammers showed inconsistent, small increases with variable discrimination. The last two sessions of baseline showed only two hammer improvisations with discrimination below 50% and dropping. Because of the declining baseline, train sufficient exemplars began in session 10,

AYSA

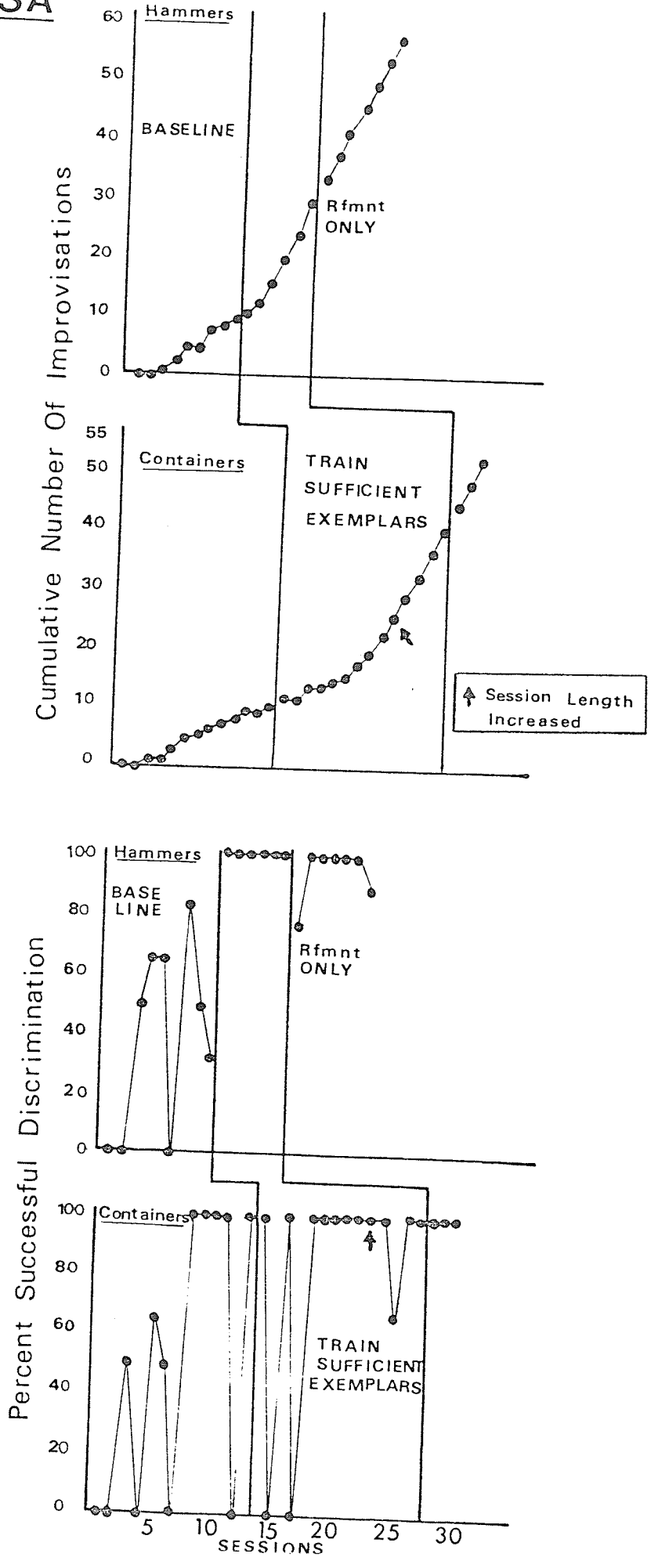


Figure 4. Cumulative number of improvisations and percent successful discrimination for Aysa in hammers and containers.

where the improvisation rate showed little change for two sessions, then rose to a maximum immediately. Discrimination rose sharply to 100% and remained there. When sufficient generalization was reached, reinforcement only began and maintained the high rate of improvisation. Percent discrimination fell slightly to 75% in the first session of reinforcement only, but then rose again to 100% for five sessions, dropping to an acceptable 80% in the final session.

Baseline improvisations of containers were consistently low, at one per session. Percent discrimination was variable at first, remaining high near the end of baseline. With so few improvisations, it was judged that no generalization of training across tasks had occurred and that even 100% discrimination was not significant with only one improvisation per session. The number of sessions used to assess across task generalization were few, but again time limitations influenced the decision to begin training. The first seven sessions of training showed little effect on either rate of improvisation or percent discrimination. Daily improvisation then began to increase steadily, reaching maximum shortly after the increase in session length. Discrimination after seven sessions of training was steady at 100% except for one session after the session length increase where it fell to 66% perhaps due to more time to make attempts. Once sufficient generalization was reached, the reinforcement only phase maintained maximum improvisation and 100% discrimination.

Sufficient Exemplars Plus Descriptive Praise.

Wendy. Figure 5 presents the cumulative improvisation and percentage discrimination data for Wendy. Baseline improvisation and discrimination were zero for containers. Training began in session five with no effect. Following a six day absence her rate of improvisation rose steadily as did percent discrimination. Within three sessions of Wendy's increased daily rate of improvisation her discrimination rose and stabilized at 100%. In session 23, session length was increased which increased rate of improvisation to the maximum. Discrimination shows a sudden drop to 75% in session 23, which probably represents an effect of more time to try out items. The same effect was noted in Jason and Aysa's data. Once sufficient generalization was reached, Reinforcement Only, maintained maximum improvisation with 100% discrimination.

Wendy's improvisation in hammers during baseline were few, with most occurring after her absence, though she did level off at zero for the three points at the end of baseline. No generalization across tasks was noted. However, containers were not yet fully under training control, but her long absence had made time limitations again a critical factor. Her discrimination data revealed two peaks at 50%, when she did improvise in baseline, but discrimination fell to zero at the end of baseline. Sufficient exemplars plus descriptive praise produced

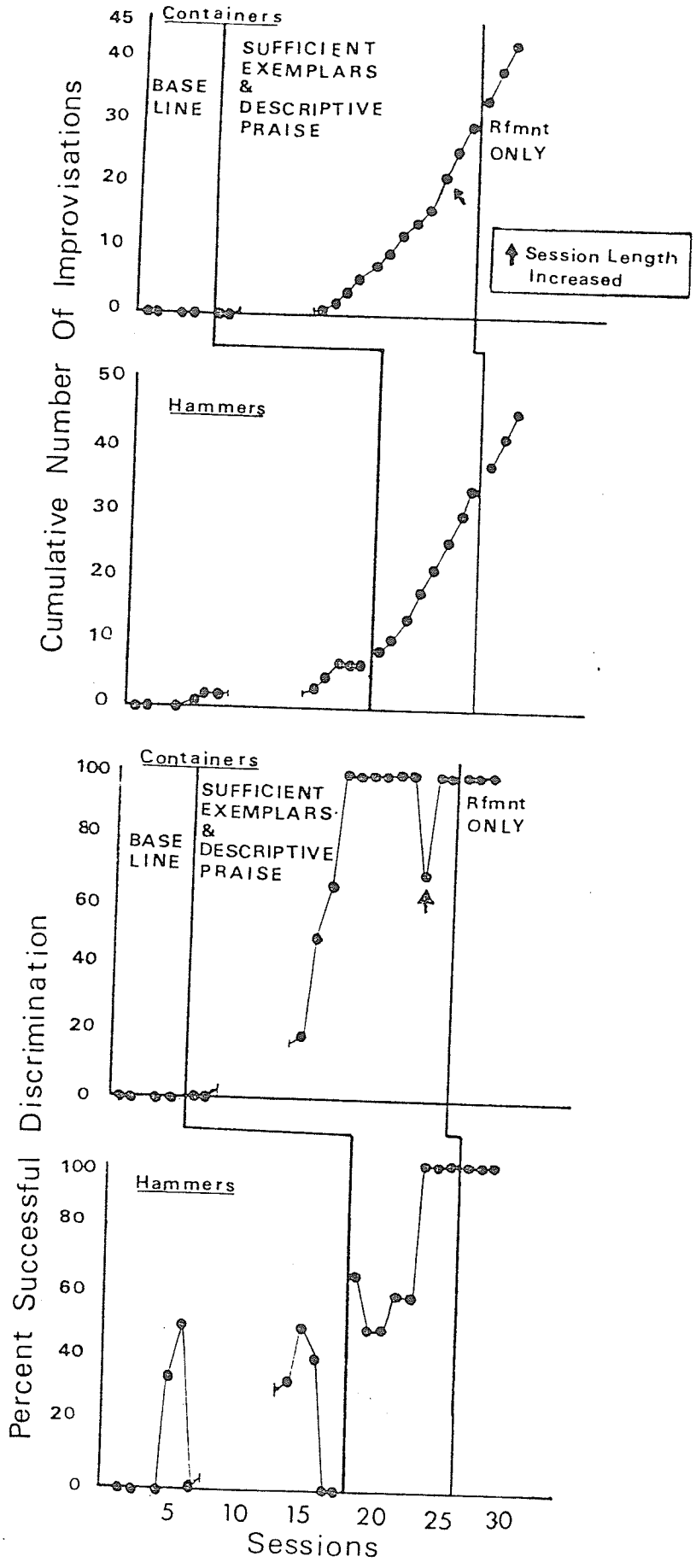


Figure 5. Cumulative number of improvisations and percent successful discrimination for Wendy in hammers and containers.

an immediate, steady rate of improvisation. Discrimination rose sharply and remained above 50% eventually reaching a stable 100% in the last three sessions of training. Once sufficient generalization was reached, maximum improvisation and 100% discrimination were maintained through three reinforcement only sessions.

Ché. Figure 6 shows the data for Ché, in cumulative number of improvisations and percent successful discrimination. In containers, Ché began with several improvisations, with discrimination generally at or below 50%. After three consecutive sessios of zero improvisation, Ché produced one improvisation with 100% discrimination. Training began at that point despite the upswing, however small it was, because Wendy's indefinite absence threatened the across subject multiple baseline design. Very little change in improvisations could be seen over baseline. However, his discrimination level rose above 50%, although it was extremely variable. In session 23 when session length was increased Ché began improvising at or near maximum; his discrimination remained variable dipping to a low of 40% before reaching sufficient generalization. Due to a lack of time and items, only one session was conducted in reinforcement only, where maximum responding and 100% discrimination were maintained.

Ché's hammer improvisation baseline showed no responding at first, then several improvisations, becoming stable at one improvisation per session. No generalization of training ef-

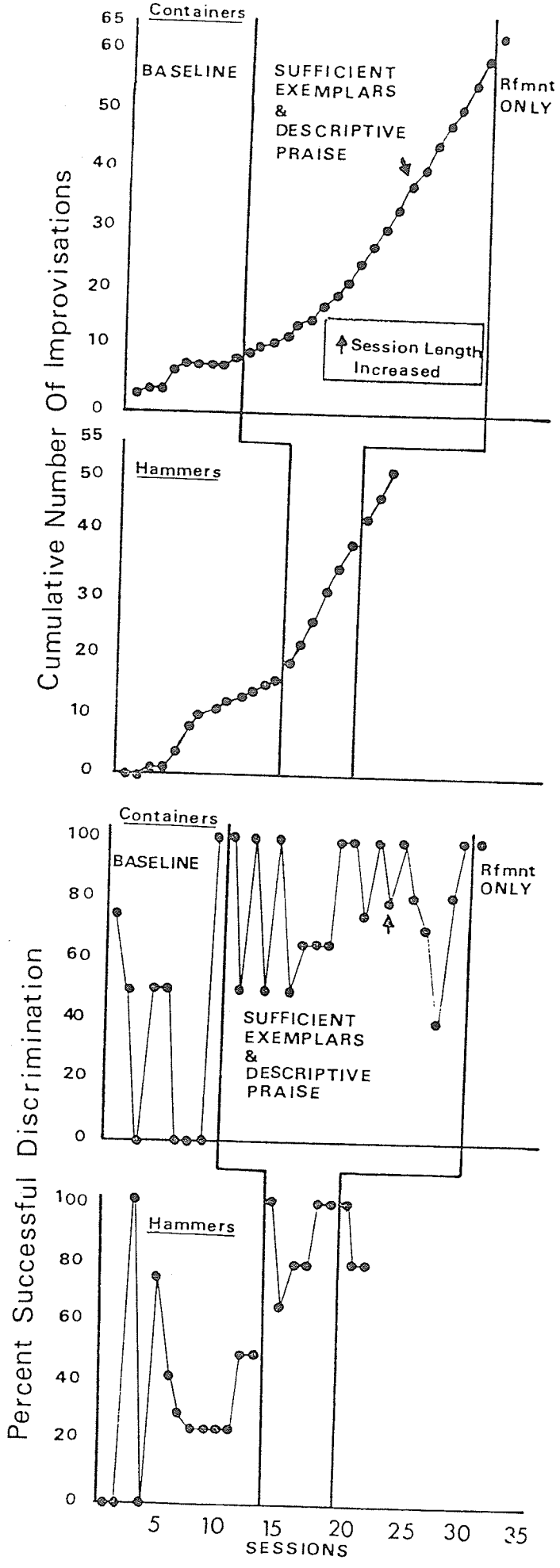


Figure 6. Cumulative number of improvisations and percent successful discrimination for Ché in hammers abd containers.

fect was seen across tasks during four sessions although minimal control had been established in containers (time limitations were again critical). Percent discrimination in baseline was highly variable at first, becoming stable at 33% and then 50%. With the low improvisation rate and discrimination level, training began in hammers. The rate of improvisation rose quickly and steadily to its maximum. Discrimination rose above 50% and continued to rise to 100%. Once sufficient generalization was achieved the rate of improvisation remained at a maximum and discrimination remained at or above 80% in Reinforcement Only sessions.

Daniela. Figure 7 presents the cumulative improvisation and percent discrimination data for Daniela. In baseline there was little improvisation, for containers becoming zero for the three sessions at the end of baseline. Her percent discrimination showed great variability although most scores remained well under 50%. The last two baseline sessions reflect zero discrimination. Training began in session 14, with no change due to training in that session. Rate of improvisation soon began to increase, however, reaching maximum just before the session length increase. Discrimination rose to 100%, then fell slightly to 75%, then rose again to 100% for two sessions. Once sufficient generalization was reached, maximum improvisation and 100% discrimination were maintained in three Reinforcement Only sessions.

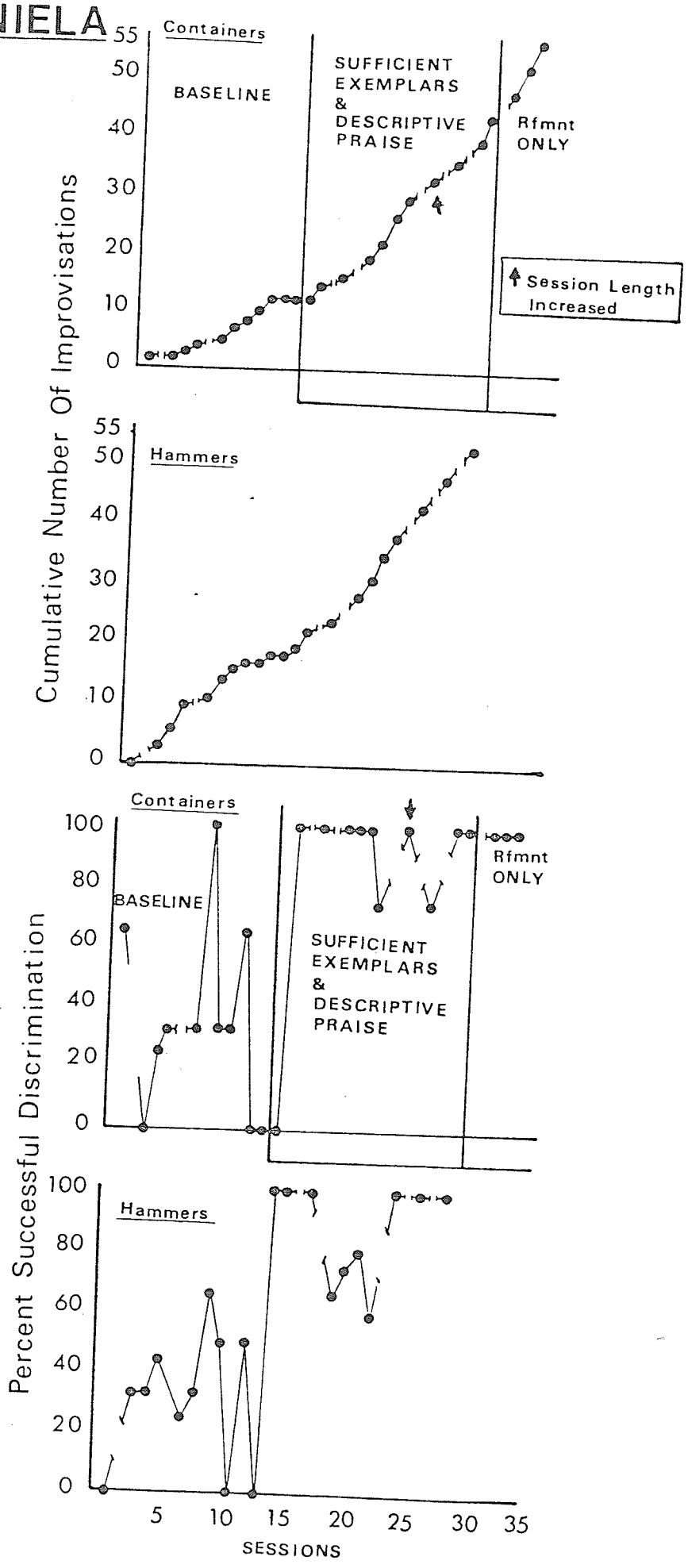


Figure 7. Cumulative number of improvisations and percent successful discrimination for Daniela in hammers and containers.

Daniela was not trained in hammer improvisation. Initially she improvised at quite a high rate, although her percent discrimination remained below 50%. Around session eight she began to improvise less and her discrimination became more variable. In session 14, when container training began, her discrimination rose dramatically to 100% and remained there for three sessions. In session 15, she began to improvise at a higher rate, becoming at or near maximum by session 20 which suggests generalization of training effects across tasks. Her discrimination level dropped for four sessions but returned to 100%. Daniela reached the criterion of sufficient generalization without training, so sessions were terminated.

Mediate Generalization

Delia. Figure 8 presents the improvisation and discrimination data for Delia. Baseline, revealed no hammer improvisation and corresponding zero discrimination. Training began in session five. Training showed a low, but steady increase in improvisations, becoming much steeper midway through training. Her discrimination data, rose considerably above zero but remained below 50%, to subsequently rise above 50% in the last half of training. There was great variability in discrimination throughout training. Once sufficient generalization was reached the maximum rate of improvisation was maintained and

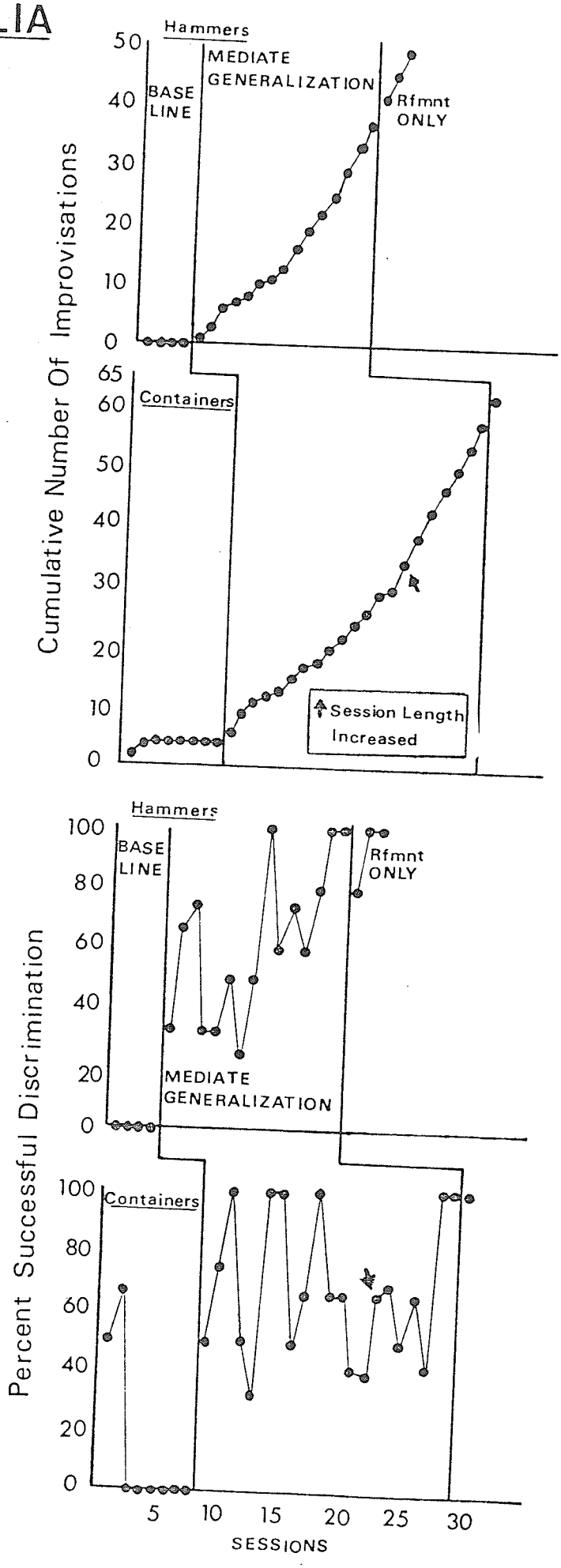


Figure 8. Cumulative number of improvisations and percent successful discrimination for Delia in hammers and containers.

discrimination remained above 80%.

Delia made four container improvisations with 50% discrimination, in the first two baseline sessions. No further improvisations were made, and no generalization of training effects across tasks was evident though sessions were limited due to time constraints. Once training began her rate of improvisation was consistent but below maximum. Discrimination was above 50% but extremely variable. When session length was increased, her improvisation rate reached maximum, but little stability was seen in discrimination. Once discrimination reached 100% and sufficient generalization was reached, the high improvisation rate was maintained in only one Reinforcement Only session. A single Reinforcement Only session was used because of a lack of tool items preventing further sessions.

Alvin. Figure 9 presents the improvisation and discrimination data for Alvin. In baseline for hammers, Alvin showed little improvisation, with variable, but low discrimination. Training had a slight effect initially, with rate of improvisation then rising quickly. Discrimination showed the same slow initial effect, but after two sessions rose to 100%, and remained above 80%. Once sufficient generalization was reached, Reinforcement Only maintained maximum improvising but discrimination fell sharply to 66%, but did recover to remain above 80%.

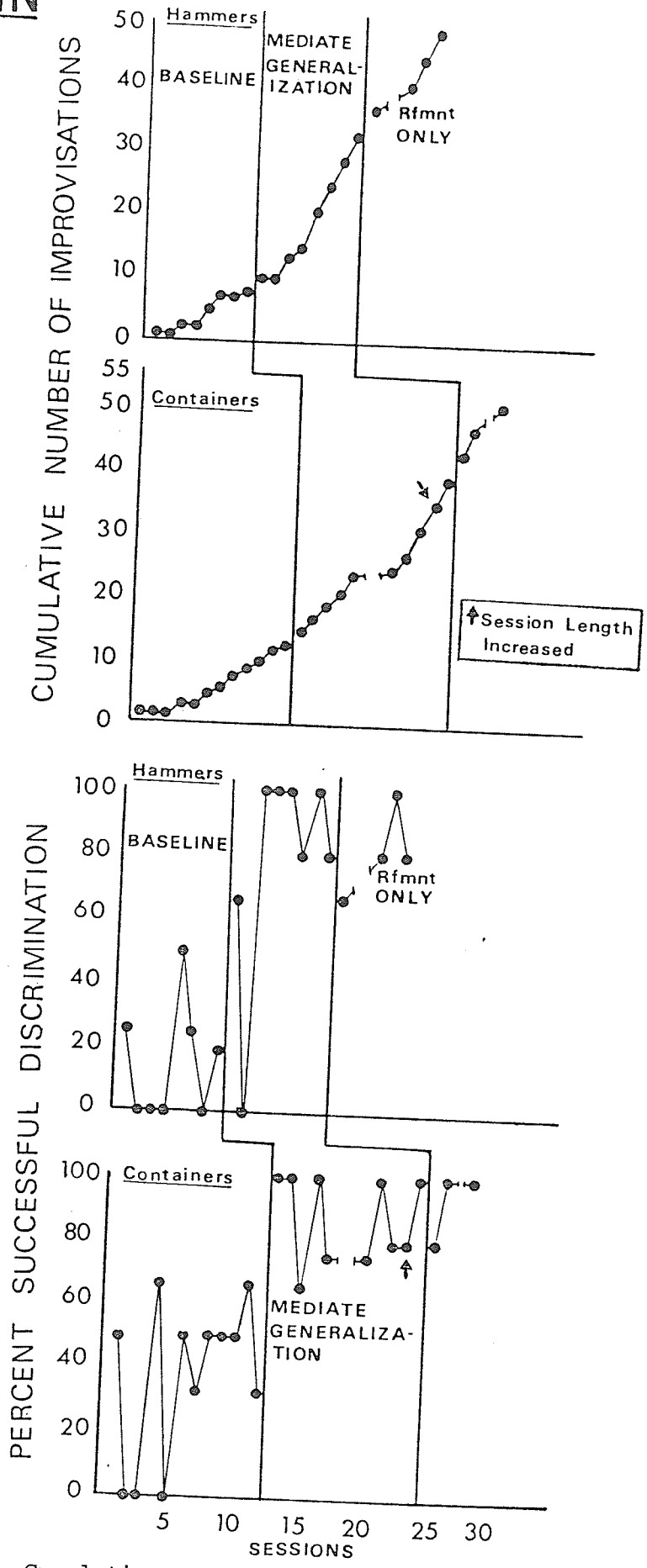


Figure 9. Cumulative number of improvisations and percent successful discrimination for Alvin in hammers and containers.

Alvin's baseline in containers showed very little improvisation for five sessions, then steady low level improvisation of one per session. No generalization of training effects across tasks was noted, although the numner of sessions to do so was limited by time constraints. Alvin's percent discrimination demonstrated a lot of variability with only two sessions above 50%. When training began it had a slight effect on rate of improvisation. Shortly after his two day absence his rate of improvisation reached its maximum. Training did produce an increase in discrimination. Although discrimination was variable, it remained above 50%. Once sufficient generalization was reached the rate of improvisation remained high and discrimination at or above 80% in Reinforcement Only sessions.

Tanya. Figure 10 presents the improvisation and discrimination data for Tanya. Baseline improvisation and discrimination in hammers was zero. Training produced very large increases in her rate of improvisation, already at a maximum in three training sessions. Her discrimination data had risen above 60% and reached 100% when she was withdrawn from the day care centre unexpectedly.

Container improvisation began with one improvisation with 100% discrimination in the first baseline session. No other improvisations were produced until session 15 when one more was produced with 50% discrimination, suggesting some brief across task generalizations may have occurred. Training produced two

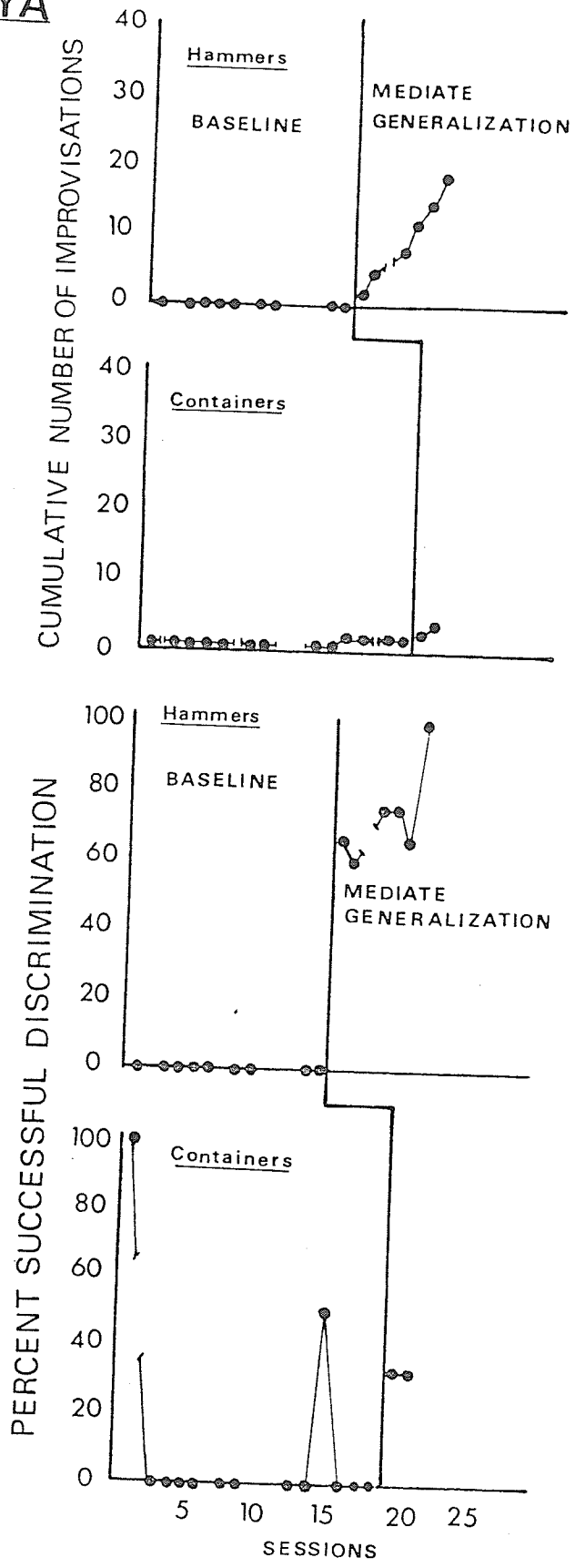


Figure 10. Cumulative number of improvisations and percent successful discrimination for Tanya in hammers and containers.

improvisations in two sessions with discrimination at a low 33%. No further training was conducted as Tanya was withdrawn from the centre.

Multiple Baseline Across Subjects Data

Train-to-Generalize. Jason and Michelle were paired in a multiple baseline across subjects design, for the train-to-generalize strategy. Figure 11 represents the cumulative improvisation and successful discrimination data for the hammer task. The results for each subject were reported in detail earlier when individual subject data was presented. The data showed that Jason produced very few hammers successfully in baseline, whereas Michelle did at first then later produced only two in four sessions. Training with descriptive social reinforcement alone had essentially no effect on rates of improvisation or percentage successful discrimination as performance did not change from baseline levels. Dramatic increases were noticed for Jason in session 19 and Michelle in session 23, after mediate generalization training had begun in containers.

Figure 12 contains Jason's and Michelle's data for the container task. Jason showed no improvisation of containers in baseline and Michelle showed only a few during a limited number of sessions. Train-to-generalize produced no change in either subject's performance. Mediate generalization

TRAIN-TO-GENERALIZE

HAMMERS

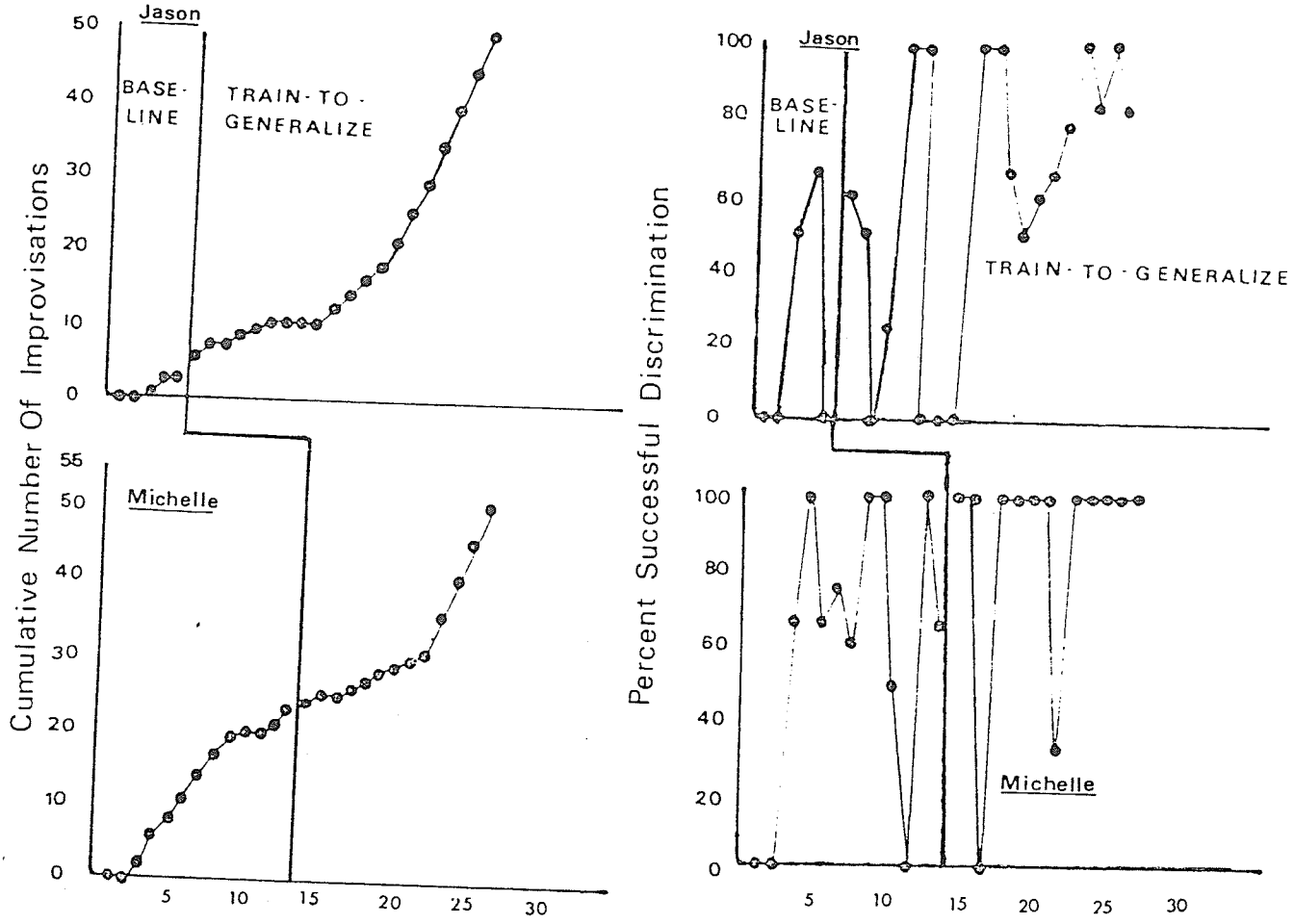


Figure 11. Cumulative number of improvisations and percent successful discrimination for Train-To-Generalize in hammers.

TRAIN-TO-GENERALIZE

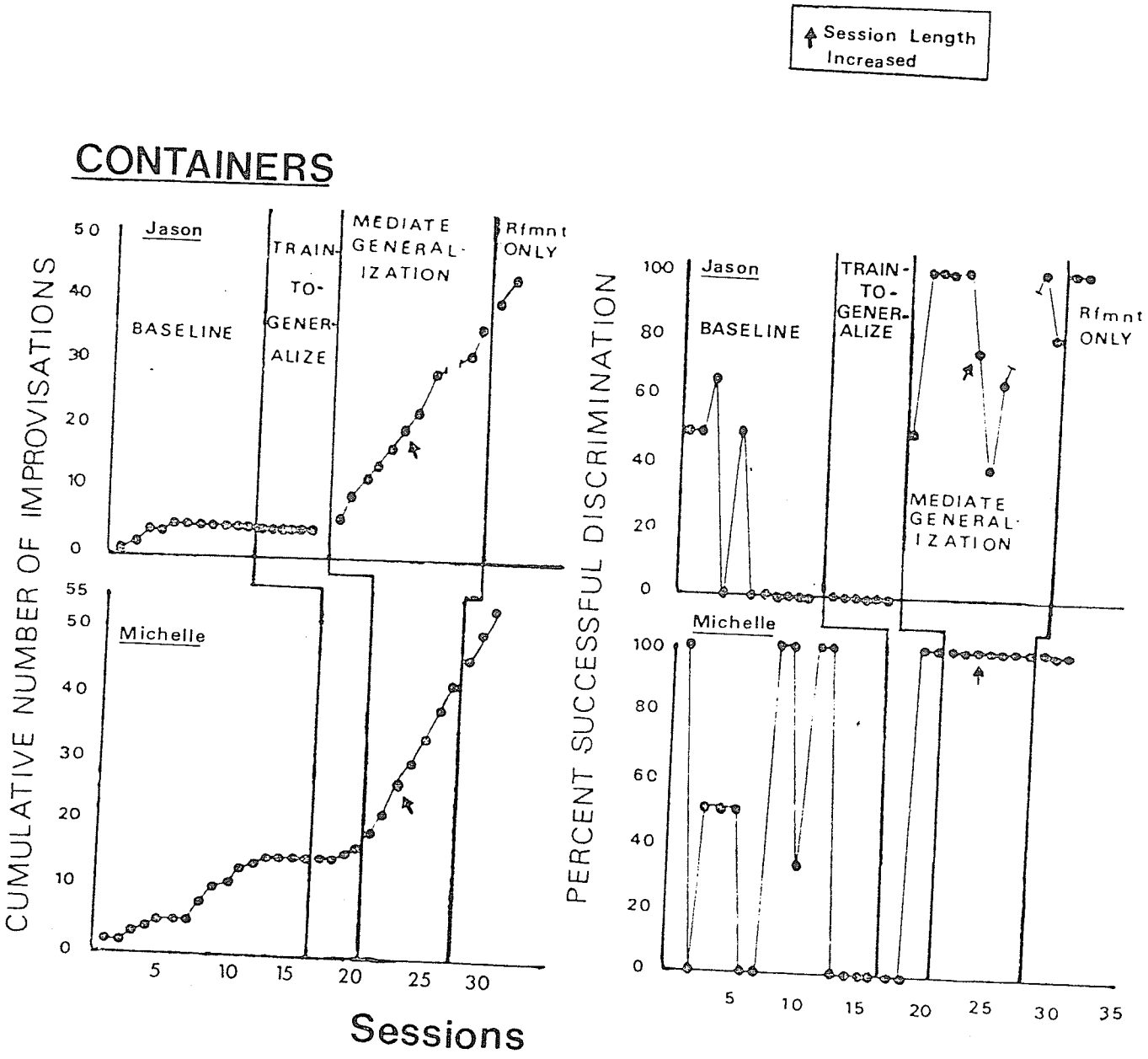


Figure 12. Cumulative number of improvisations and percent successful discrimination for Train-To-Generalize in containers.

training produced large, rapid increases in cumulative improvisations as well as increasing discrimination to levels consistently higher than baseline. Both children showed very stable, high level discrimination and improvisation in the Reinforcement Only phase.

Train Sufficient Exemplars. Sylvia and Aysa were paired in a multiple baseline across subjects design for the train sufficient exemplars strategy. Figure 13 presents the cumulative improvisation and successful discrimination data for Sylvia and Aysa in the hammer task. Sequential application of exemplars alone after baseline, produced gradual increases in cumulative improvisations and successful discrimination for Sylvia and rather more dramatic and immediate increases in both dependent measures for Aysa. Both children maintained high levels of improvisation and discrimination in the Reinforcement Only phase.

Figure 14 represents the cumulative improvisation and successful discrimination data for the container task. Both children displayed very gradual improvisation with low and/or variable discrimination in baseline. Sequential application of training produced a slight increase in improvisation initially and increased discrimination substantially. Both children showed larger increases in rate of improvisation after the session length was increased by five minutes. After reaching sufficient generalization, both maintained high levels of improvisation and 100% discrimination for three sessions.

TRAIN SUFFICIENT EXEMPLARS

HAMMERS

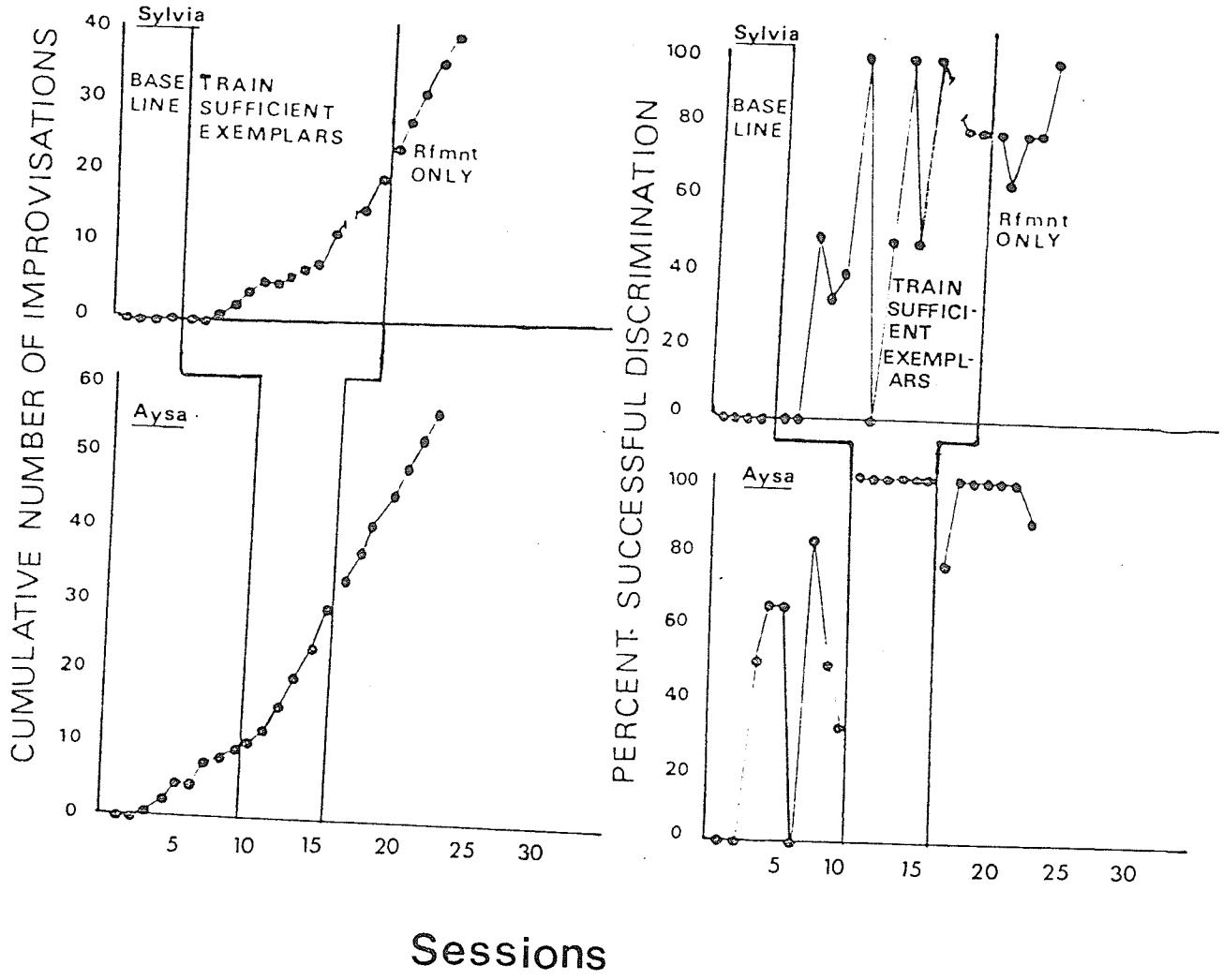


Figure 13. Cumulative number of improvisations and percent successful discrimination for Train Sufficient Exemplars in hammers.

TRAIN SUFFICIENT EXEMPLARS

CONTAINERS

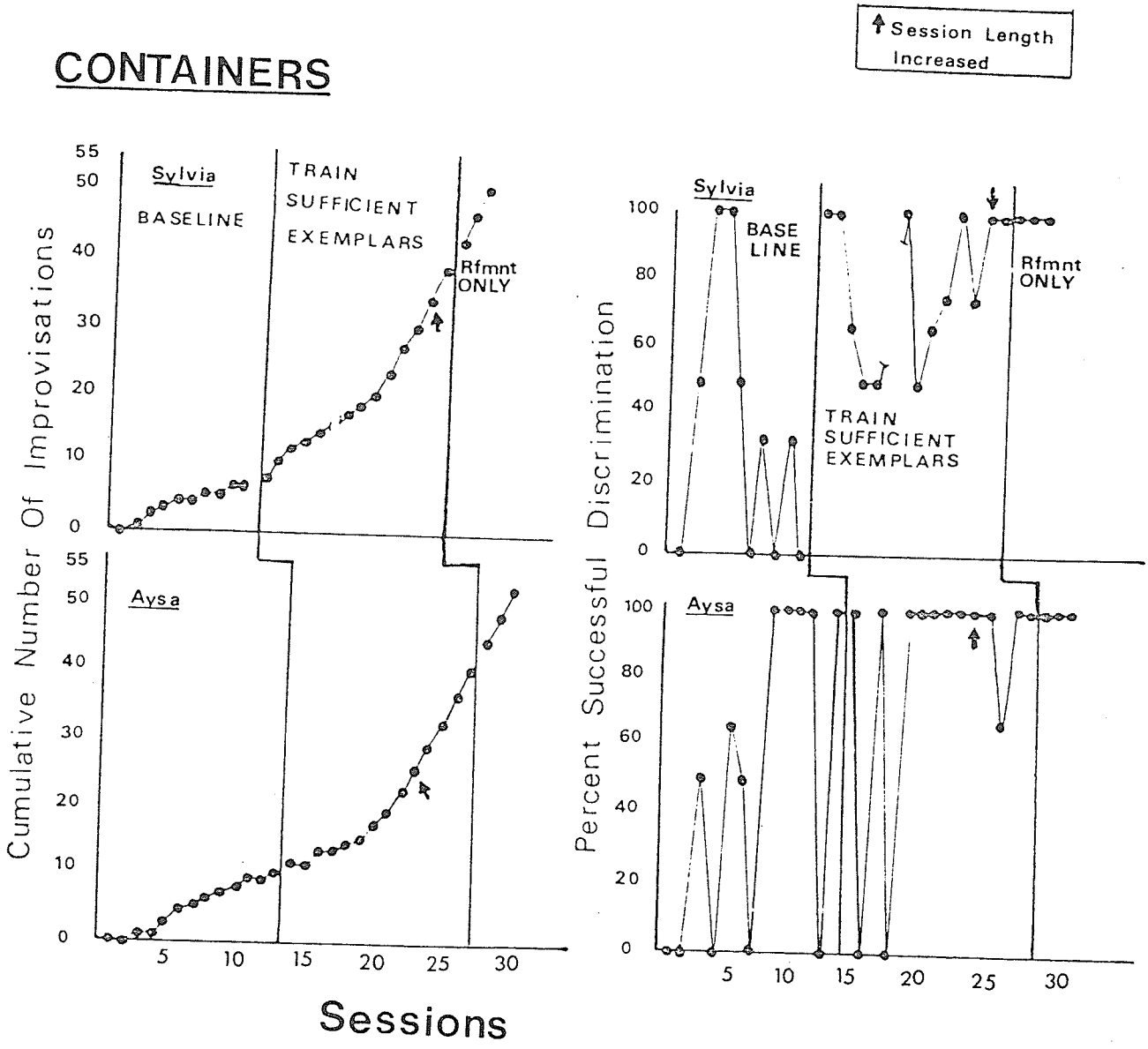


Figure 14. Cumulative number of improvisations and percent successful discrimination for Train Sufficient Exemplars in containers.

Sufficient Exemplars Plus Descriptive Praise. Ché, Wendy and Daniela were grouped in a multiple baseline across subjects design for the sufficient exemplars plus descriptive praise strategy. Figure 15 represents their cumulative improvisation and successful discrimination data for the hammer task. Training was sequentially applied to Ché and Wendy. Baseline improvisations for both children revealed a very low daily rate of improvisation with low and variable discrimination. Training produced quick, high rates of improvisation with high discrimination for both children. Once sufficient generalization was reached both maintained maximum rate of improvisation with high discrimination. Daniela did not receive training in hammers. Initially in baseline she had a higher rate of improvisation (compared to Ché or Wendy), but she also had a low and variable level of discrimination. Her cumulative improvisations and discrimination increased sharply in session 15 when training was started in containers, indicating some generalization of training effects across tasks.

Figure 16 represents the cumulative improvisation and successful discrimination data for containers. All three children showed very low cumulative improvisation in baseline with low and/or variable discrimination levels. Sequential application of training increased the rates of improvisation sharply, also raising discrimination levels reducing variability (to zero for Wendy), and maintaining maximum rate of improvisation and 100% discrimination in Reinforcement Only sessions.

SUFFICIENT EXEMPLARS & DESCRIPTIVE PRAISE

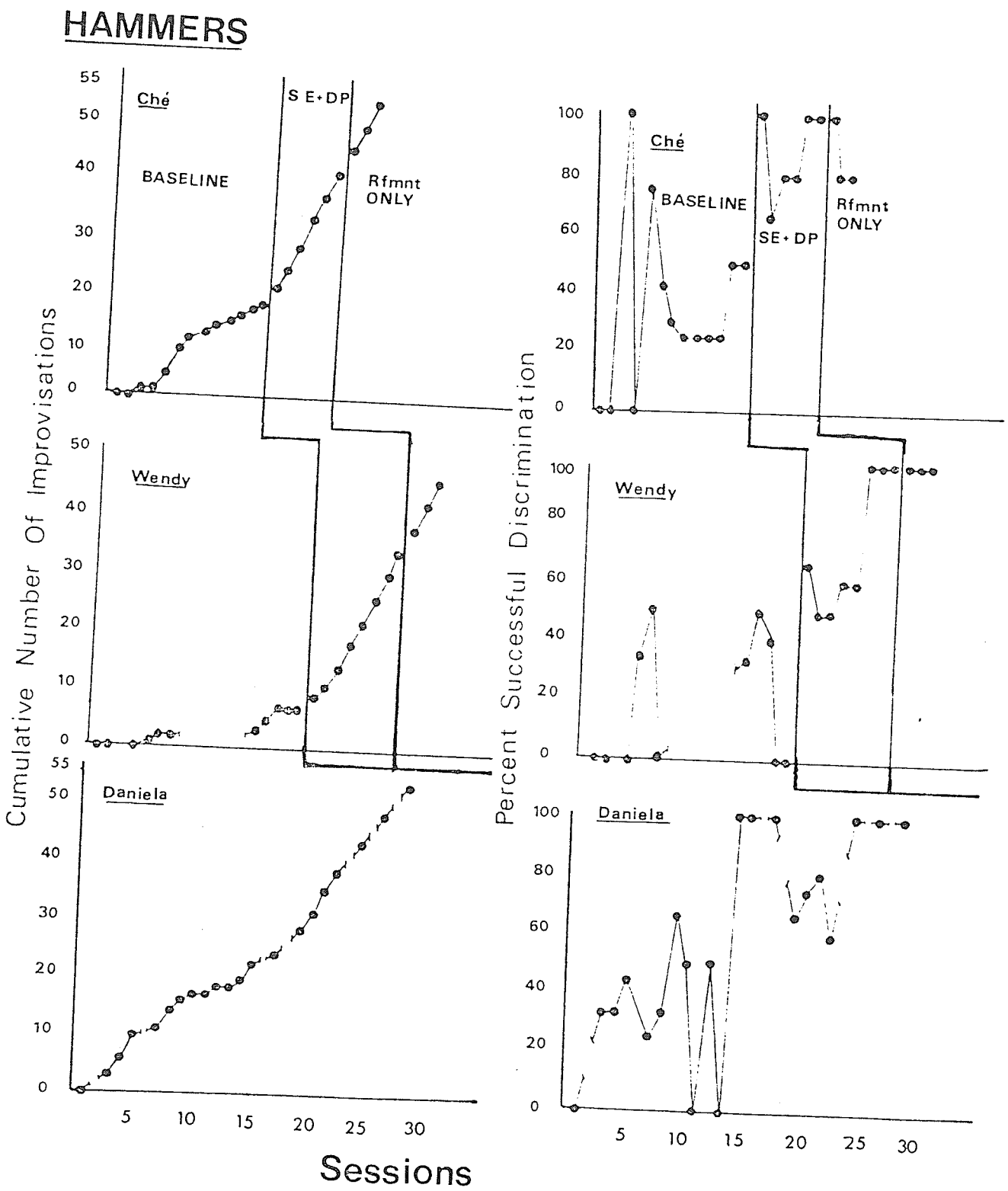


Figure 15. Cumulative number of improvisations and percent successful discrimination for Sufficient Exemplars Plus Descriptive Praise in hammers.

SUFFICIENT EXEMPLARS & DESCRIPTIVE PRAISE

CONTAINERS

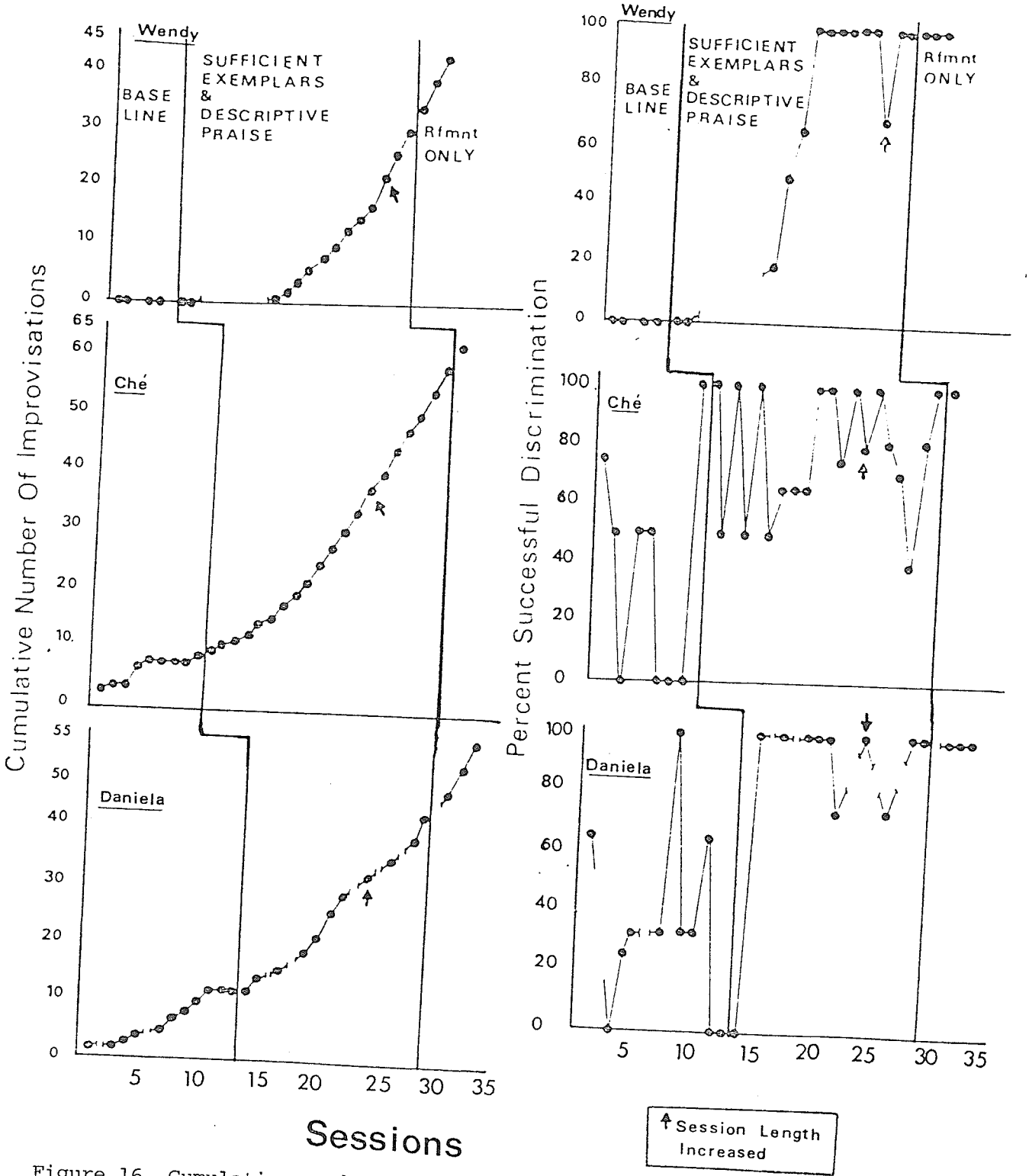


Figure 16. Cumulative number of improvisations and percent successful discrimination for Sufficient Exemplars Plus Descriptive Praise in containers.

Mediate Generalization. Delia, Alvin and Tanya were grouped in a multiple baseline across subjects design for the mediate generalization strategy. Figure 17 represents the cumulative improvisation and successful discrimination data for the hammer task. Delia and Tanya showed no improvisation and zero discrimination in baseline, while Alvin showed very few improvisations with very poor discrimination. Training produced large gains over baseline levels of improvisation with consistently higher levels of improvisation and consistently higher levels of discrimination, which were maintained in Reinforcement Only sessions.

Figure 18 represents the cumulative improvisations and successful discrimination data for containers. Delia and Tanya showed almost no improvisations in baseline with essentially no discrimination. Alvin showed a low, gradual rate of improvisation in baseline with moderate discrimination. Training increased cumulative improvisations in the children, especially once the session length for containers was increased. Training also produced consistently higher levels of discrimination in the children. Reinforcement Only sessions indicated continued high improvisation and discrimination. Tanya's training was incomplete because she was withdrawn from the day care centre.

Strategy Comparison Data.

Table 3 presents a summary of the total training sessions required to achieve sufficient generalization as well as the

MEDIATE GENERALIZATION

HAMMERS

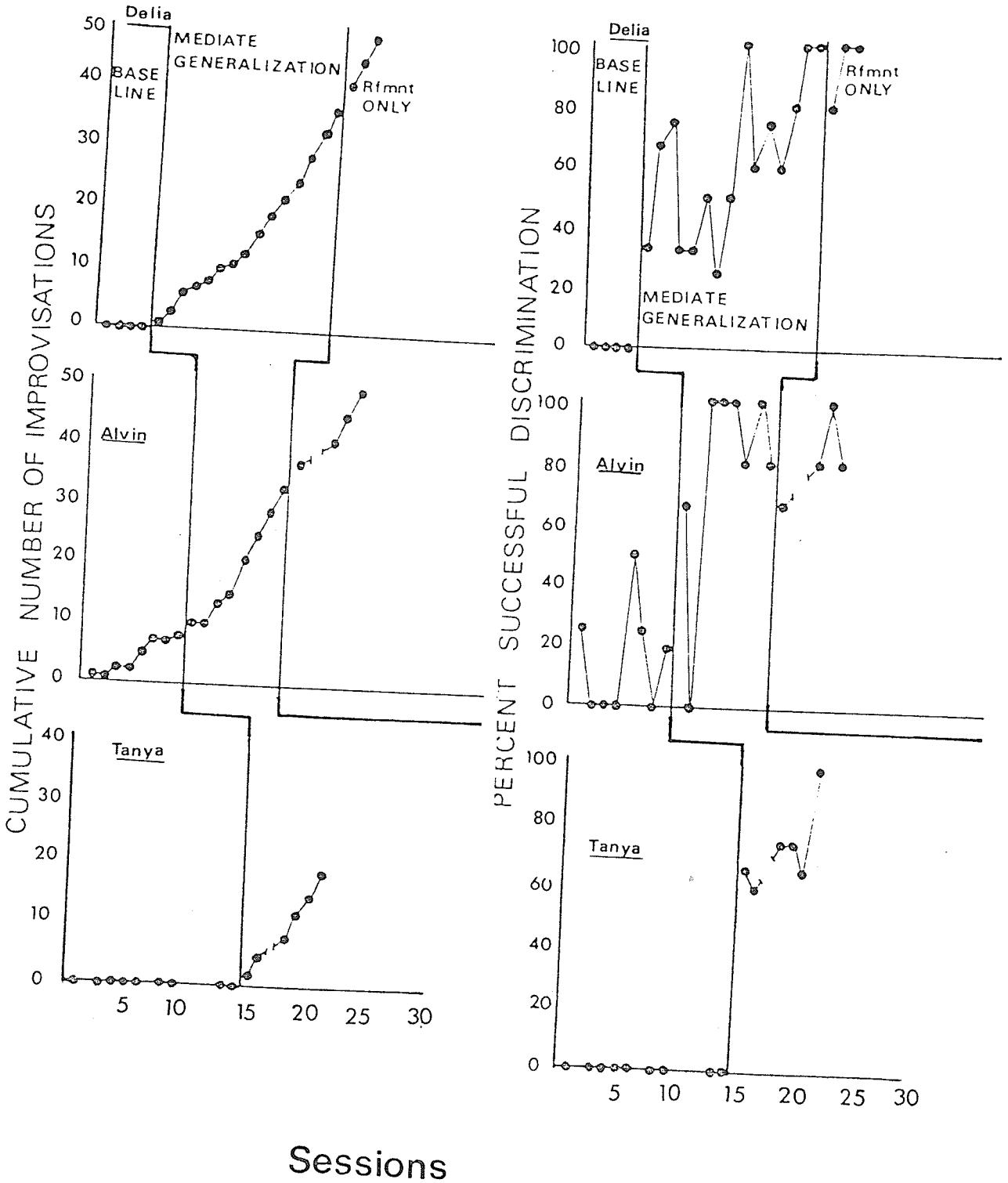


Figure 17. Cumulative number of improvisations and percent successful discrimination for Mediate Generalization in hammers.

MEDIATE GENERALIZATION

CONTAINERS

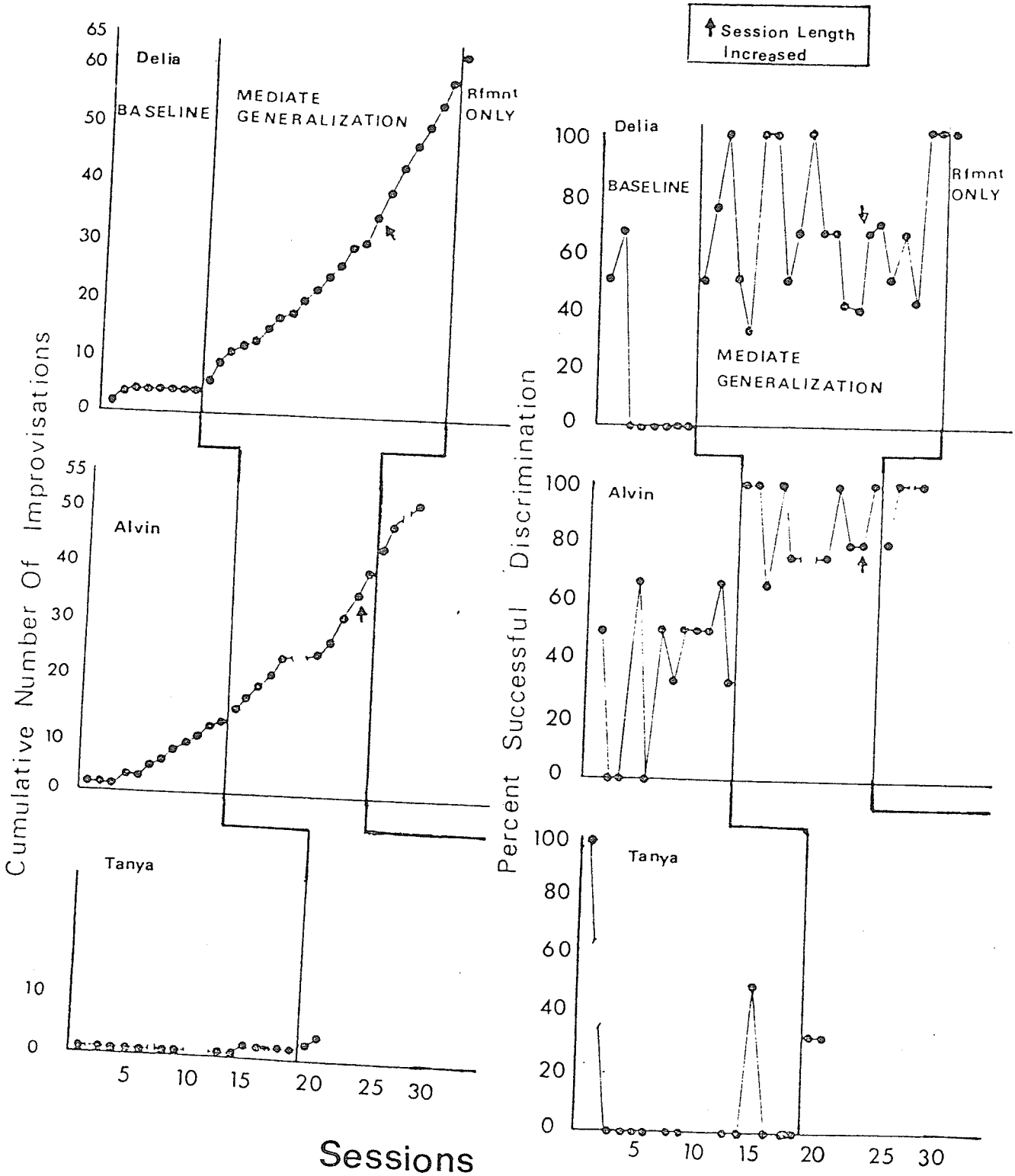


Figure 18. Cumulative number of improvisations and percent successful discrimination for Mediate Generalization in containers.

Table 3

Summary of Frequency of New Improvisations in Training

Strategy	Subject	Number of Sessions		Total Improvisations		Rate	
				Task			
		H*	C*	H	C	H	C
Train-to-Generalize	Jason	12	6	47	0	1.41	0
	Michelle	7	4	28	2	0.86	0.50
Train Sufficient Exemplars	Sylvia	13	13	20	32	1.53	2.46
	Aysa	6	14	19	30	3.16	2.14
Sufficient Exemplars Plus Descriptive Praise	Wendy	8	14	27	30	3.37	2.14
	Che	6	20	23	50	3.83	2.50
	Daniela	-	11	-	32	-	2.91
Mediate Generalization	Delia	15	21	37	54	2.46	2.57
	Alvin	8	10	24	26	3.00	2.60
	Tanya	not trained to criterion in tasks					
	Jason	-	10	-	32	-	3.20
	Michelle	-	7	-	25	-	3.57

- Task not trained.

* H; Hammers

C; Containers

total improvisations and rate of improvisation during training for each of the 10 children. The mean rates of improvisation during training revealed that train-to-generalize had the lowest rate for containers and second lowest for hammers. Sufficient exemplars plus descriptive praise produced the highest mean rate of improvisation of hammers, with train sufficient exemplars and mediate generalization about equal. The three strategies including exemplars produced approximately the same rate of improvisation in containers.

Table 4 presents the number of sessions of training, total improvisations and rates of improvisation during training from Parsonson and Baer (1978). Comparison of the data in Tables 3 and 4 revealed that the children in the present study, in all strategies, produced many more improvisations in fewer sessions with exceptionally higher rates of improvisations in both hammers and containers. The data in Table 3 indicate little difference in rates for either task, even in the strategy (Sufficient Exemplars plus Descriptive Praise), where containers was trained first and hammers second, indicating further that no order effects were present. Also, the difference in rate of improvisation between containers and hammers was not substantially different in any of the training strategies (except train-to-generalize), indicating that neither task was significantly easier to train, with hammers being trained somewhat more easily in some cases. Since the number of training sessions equals the number of exemplars

Table 4

Summary of Frequency of New Improvisations in Training Subjects in Parsonson and Baer (1978)

Subject	Number of Sessions		Total Improvisations		Rate	
	H*	C*	H* Task	C*	H*	C*
Elsie	20	-	8	-	.40	-
Jerry	28	18	6	2	.20	.10
Jimmy	18	-	4	-	.20	-
Luther	38	31	6	6	.16	.19
Ronnie	18	19	7	8	.38	.42

- Task not trained

*H; Hammers

*C; Containers

(because only one exemplar was trained per session) the number of "sufficient" exemplars for each child can also be gleaned from Table 3. There was very little difference in the number of exemplars used in hammers and containers for each of the children within and across strategies (ranging from 6 to 15 for hammers and from 10 to 21 for containers).

Discussion

The results for each generalization programming strategy will be discussed first, followed by a general discussion of problem solving and the structure of problem situations. The relationship of discrimination and generalization in creative problem solving will then be summarized. Next, major methodological issues will be approached with an eye to future research. A general conclusion appears at the end.

Train-to-Generalize

The results indicated that train-to-generalize alone, did not produce significant amounts of generalized improvisation in the container task. Furthermore most improvisations in hammers occurred only after extra training (mediate generalization) was started in containers. The data suggest that the level of the problem solving operant in the subjects was too low. Therefore, as predicted, improvisation was not appreciably influenced by the descriptive social reinforcement, since improvisation responses would have to occur in order to be subject to reinforcement control.

The initial effect of train-to-generalize in the hammer task was to set a rate of one improvisation per session; through an adventitious differential reinforcement of low rates schedule. That is, the children typically would produce one improvisation, would then receive the descriptive social reinforcement, and would then indicate that they could not find any more hammers which terminated the session. The termination of the session without further prompting the children to find a new tool could have been adventitiously reinforcing the production of only one response followed by the statement that no further responses were possible. The children might have been reinforced because they could go play outside again which may have been more reinforcing than the session activities or might have seen the experimenter's termination of the session as approval that, indeed the children were right and no further responses were possible. Therefore, because of the low operant level, and the adventitious reinforcement schedule, the children might not have distinguished what response was desired; novel responses, not a solitary response.

Train Sufficient Exemplars

The train sufficient exemplars strategy was more effective in producing generalized improvisation within both tasks than train-to-generalize alone. The rate of improvisation in

both tasks was higher than for train-to-generalize. The immediacy and size of the effects of training exemplars in this strategy were larger than expected. The lack of effect of differential social reinforcement in the train-to-generalize tasks and the effect here without it, suggests that the children may have been reinforced by the completion of the task with a new tool. The media trap effect mentioned in Parsonson and Baer (1978, p. 379), i.e., any "material, activity or medium that eventually will maintain appreciable rates of child behavior with it, independent of any extrinsic reinforcement" could have been the reason for the amount of generalized improvisation developed in both tasks. In Parsonson and Baer (1978) the media trap was cited as a possible hinderance to success by producing perseveration with a favorite item. However, in the present study, because of the structure of the problem situation, no perseveration occurred in conjunction with the media trap. That is, if the children were reinforced by the completion of the task with an item, that same response was never possible again in a session or subsequent sessions, because the solution tool was removed from the item pool.

The proper implementation of a media trap seems highly desirable, because both Skinner (1968) and Bijou (1976) have indicated that creativity training procedures must eventually program the transfer of behavioral control from the teacher's reinforcement to natural contingencies, i.e., reinforcement

of solving problems. Proper use of a media trap presumably reflects the development of conditioned reinforcers which are part of the task (for which less analytic approaches have often invoked the circular label "intrinsic reinforcement"). Children, of course, usually have substantial prior histories involving the use of hammers and containers making the beneficial effects of a media trap possible. However, it would be possible to develop a task that would become a conditioned reinforcer if such a task were not already present in a subject's repertoire, simply through differential reinforcement.

Sufficient Exemplars Plus Descriptive Praise

The strategy combining train-to-generalize and train sufficient exemplars produced substantial generalized improvisation within each task, with a slightly higher rate of hammer improvisation than any other strategy. The exemplar training fostered novel improvisations, and their occurrence was strengthened by the descriptive social reinforcement. Some generalization across tasks, from containers to hammers was demonstrated in Daniela's data.

This combination of procedures was the same as in Parsonson and Baer's (1978) research. As a group in Parsonson and Baer's study, three children in the hammer task produced 16 improvisations and five children in the container task produced 31 improvisations. In the present study, two children in the hammer task

produced 74 improvisations and three children in the container task produced 139 improvisations. Three less children in the present study produced 166 more improvisations. The mean rates of improvisation per session during training were substantially lower in Parsonson and Baer's study than here (cf., Table 3 & 4). The staggering difference in amount of generalized improvisation between the two studies would seem directly due to the more systematic problem situation maintained in the present study, by not allowing old solutions to be available. As in the train sufficient exemplar results a beneficial "media trap" would have resulted. The structure of the problem in this study was such that the extrinsic experimenter reinforcement was delivered on a novel for this session and a novel for this experiment basis, because the same response could not occur twice. This avoided perseveration by the children which would have resulted in extinction from experimenter reinforcement. It was clearer here, that the abstraction of novel task completion was being reinforced not just task completion with a particular item. Thus in this manner, if a media trap occurred, it would be working in cooperation with the extrinsic social reinforcement of the experimenter.

Other reaons for the massive difference in results are that the children were not reinforced for coming to the sessions or for using the exemplar items after the experimenter modeled them. These procedures helped to make the nature of

the desired responses clear to the children, i.e., completing the task with a new tool. The target response was further facilitated in that the exemplars came from the tray of items and frequently had been on the tray in previous sessions. Thus children were taught to search the tray for an item, pick it up and complete the task, not just to take an item from the experimenter and to complete the task as in Parsonson and Baer's (1978) situation. Finally, the fact that the improvisation items were all simple improvisations (i.e., they were concrete items which functioned as the desired tool without assembly) and were more numerous than in Parsonson and Baer's (1978) case (i.e., 24 hammers; 14 containers vs. 53 hammers; 85 containers) enabled the children to be more successful quickly and more easily. Initial success and reinforcement would then establish stimulus control early in training.

Mediate Generalization

The results of the mediate generalization training indicated that verbal cues, exemplars and descriptive social reinforcement did facilitate tool improvisation. The highest rate of improvisation for containers was recorded by Michelle during mediate generalization training. Michelle also demonstrated consistently perfect discrimination. Rates of improvisation in other tasks and children were about the same as train sufficient exemplars and sufficient exemplars plus descriptive praise. However, when Jason and Michelle were trained in

container improvisation using mediate generalization, both showed, sequentially, clear and consistent generalization of that training to the hammer task not receiving mediate generalization training. Also, it can be anecdotally reported that the children in this strategy were able to say what they were looking for (i.e., "something hard"; "something big"). Bijou (1976) has indicated that society values two types of knowledge: (a) knowledge of how to do things (e.g., how to go about solving a missing hammer problem); and (b) knowledge about things (e.g., hammers are hard). While all exemplar strategies used in this study successfully trained improvisation of tools, only the children trained with mediate generalization could verbally report what they were doing. However, anecdotal reports are only suggestive, therefore intense research should be conducted to assess the extent of the children's knowledge about things.

Mediate generalization logically then is an ideal strategy to use to further assess how verbal behavior can function as a generalization programming strategy and in problem solving. Future work should attempt to record all of the children's verbal responses and provide reinforcement for accurate verbal responses in assessment of generalization in order to identify the most salient verbal stimuli. Different verbal responses could then be tested for their ability to promote generalization of a trained problem solving process.

Finally only in mediate generalization training was there a substantial amount of generalization of training effects across tasks. Many of the other assessments of across task generalization for the other strategies were severely limited by time constraints; however, the across task generalization produced by mediate generalization was clear and occurred soon after training was started in the other task. Mediate generalization has demonstrated great potential to develop both knowledge of how to do things and knowledge about things and to produce generalization of these behaviors to other problem solving contexts.

Problem Situations and Problem Solving

Skinner's (1953) analysis of problems and problem-solving revealed: (a) that a problem exists when a subject is being prevented from making a learned response; (b) that no previously used solution is possible; and (c) that there is an important distinction between the final solution and the problem solving processes which lead to the solution. The present study has verified Skinner's analysis in a number of ways by using techniques for the development of operant behaviors and generalization.

All of the children in this study were able to perform the tasks chosen (i.e., they could pound pegs and store marbles), thus fulfilling condition (a). No previously used

solution was ever present therefore the children were always faced with a new problem, thus fulfilling condition (b). Finally, all of the children were taught to look at the tray, pick up an item and complete the task, thus fulfilling condition (c). This study has demonstrated the importance of the whole problem setting in order to facilitate accurate, efficient problem-solving.

Discrimination Training and Generalization Programming

It would appear that the problem-solving operant is a combination of discrimination and then generalization for the problem solver. First, the subject must discriminate between essential features of a solution and non-essential features, then respond along the generalization gradient representing solutions available. Further research in making problem solving training more efficient should then identify the underlying generalization gradient of potential tool and exemplar items. A clear assessment of the difficulty in discrimination of the essential features for each item would aid in the organization of a constant level of problem difficulty and in the best presentation of exemplars.

It is clear from the results of this study that a low operant level can be a crucial consideration in problem-solving training. Without a well established discrimination of correct responses, a subject cannot generalize to perform other correct

responses. Modeling exemplars of the whole process is extremely important to teach discrimination of correct responses from incorrect responses. Proper use of generalization programming (i.e., exemplar training and mediate generalization) has been shown in the present study to develop accurate discrimination as well as to foster generalized responding.

Mediate generalization training was shown to be a powerful strategy to facilitate correct discrimination, generalized responding, generalization to other tasks and accurate verbal descriptions of the solution by the subjects. It would seem to be a very potent generalization programming strategy because of its versatility in problem-solving training.

This study did not systematically research the best combination of diverse exemplar presentation. Nor did it clearly establish the underlying generalization gradient in the exemplars and tool items. However, in the light of results where the structure of the problem situation itself was so critical, the establishment of the best presentation of exemplars and their location on a generalization gradient seems extremely important to maximize generalization.

Methodology and Future Research

The results obtained in this study suggest that generalization programming research must clearly describe all aspects of a strategy's procedures (as was done in the present study), to

facilitate future expansion of each techniques' potent features. Furthermore the importance of proper implementation of the train sufficient exemplars strategy, has been emphasized. That is, exemplars should model the desired response as accurately as possible, e.g., exemplar items taken from the same tray as the child's potential tool items and not from a separate tray as in Parsonson and Baer's (1978) procedure.

This study overwhelmingly demonstrated the need for multiple measures of creative problem solving behaviors. Measuring cumulative number of improvisations in conjunction with the successful discrimination measure indicated how much genuine improvisation was established. In this way the procedures which are no better than trial-and-error means can be uncovered. Prolific responding alone may appear impressive on a single measure, but once observed in conjunction with a discrimination measure of high accuracy, it becomes a valuable well established skill. More measures of creative problem solving would be of value to further the study of generalization across tasks. The most salient verbal labels would aid in training generalized discrimination across problems. A measure of actual training time would aid in assessing the most expedient training strategies. All these measures and more would make the establishment of a criterion of sufficient generalization easier.

It is imperative that a criterion of sufficient general-

ization be established in every particular case, to foster comparison of various generalization programming strategies on the same and different tasks. Comparison of techniques would aid in establishing the most expeditious procedures for teaching complex behaviors in specific subject areas. Without further specification of "sufficient" generalization and "sufficient" exemplars the concept itself will remain mysterious and elusive.

All applied research is vulnerable to difficulties in controlling all aspects of a subject's environment and of choosing the best measure of a particular behavior. A major problem in this study arises from a lack of clear behavioral control (i.e., on both cumulative number of improvisations and percent successful discrimination) being established in one task before sequentially beginning training in another task. This problem was essentially produced by the lack of an indefinite number of tool items and subject time. It was uncertain, too, how long a child could be maintained in baseline without developing an aversion to the task by consistent failure. Also because it was not possible to control all aspects of the children's environment, thus play outside could have been more reinforcing than play with the experimenter. It would be possible to establish more rigid control over subject availability and competing reinforcers, but it is not always possible in an applied setting such as a day care centre or classroom.

The problem of early phase changes does not effect the interpretation of results within each task. However, because complete control in the first task was often not achieved when training on the second task was started does limit the interpretation of the results of each strategy's ability to produce across task generalization. Future research should where possible avoid subject time constraints, so as to establish clear behavioral control in one task to observe each strategy's ability to produce across task generalization.

Training was also initiated in three clear instances where the change from baseline is ambiguous (Michelle-hammers; Aysa-containers; and Ché-containers). Again this was partly due to subject time constraints but also to a lack of a pre-established stability criterion. The ambiguity of these cases does not alter the conclusions because clear shifts are made in other tasks for these children and for other children in the strategies. Future research should set a criterion of stability before beginning data collection to ensure that all cases would be clear demonstrations of control.

Conclusion

The development of tool improvisation in the present study was clear and massive. Every subject in each strategy (except train-to-generalize) produced many times more successful tools in both tasks after training. The results are

sequentially replicated across 10 children in two tasks each, in the four generalization programming strategies.

The elaborate application of the four strategies and the attention given to the details of the problem situation worked in cooperation with a task media trap effect, to further the achievements of generalization programming by making the reinforcement contingencies a part of the problem-solving behavior itself. "Generalizing" from these results, it becomes possible for parents and teachers to identify and arrange conditions and contingencies to increase the probability of successful problem-solving and originality. While additional research is still needed, to develop the most expeditious procedures for teaching creative behaviors, the benefits of behavioral analysis are mounting.

As more tasks are trained in a single individual or individuals, in music, art, mechanical technology or mathematics, perhaps more and more consistent generalization across untrained tasks will be seen. Training which establishes creative behavior in just one task is valuable to the individual and society. However, mediate generalization appears to be a potent combination of valuable procedures with the clear potential of developing creative behavior not only within a task but also across other untrained tasks. Whatever mediate generalization's eventual contribution, it is readily apparent that "creativity" is yielding to a behavioral analysis. Only then will less mystery surround those individuals who are creative and will more reinforcement be given to more people.

Reference Notes

1. Goetz, E. M., A review of the training of creativity as an operant behavior. Paper presented at the Midwestern Association for Behavior Analysis convention, Chicago, May, 1977.

2. Parsonson, B. S. & Baer, D. M., Appendices to Parsonson, B. S. and Baer, D. M., Training generalized improvisation of tools by preschool children, 1978. Available from B. S. Parsonson, Psychology Department, University of Waikato, Private Bag, Hamilton, New Zealand.

References

Ballard, K. D. & Glynn, T., Behavioral self-management in story writing with elementary school children. Journal of Applied Behavior Analysis, 1975, 8. 387-398.

Baer, D. M., Peterson, R. F. & 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.

Baer, D. M., Wolf, M. M., & Risley, T. R., Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1968, 1, 91-97.

Bijou, S. W., Child development: the basic stage of early childhood. Englewood Cliffs, New Jersey, Prentice-Hall, 1976.

Bringhan, T. A., Graubard, P. S. & Stans, A., Analysis of effects of sequential reinforcement contingencies on aspects of composition. Journal of Applied Behavior Analysis, 1972, 5, 421-429.

Cole, H. P. & Parsons, D. E., The Williams Total Creativity Program. Journal of Creative Behavior, 1974, 8, 187-207.

Ferster, C. B., Culbertson, S. & Perrott-Boren, M. C., Behavior Principles, 2nd Edition, Englewood Cliffs, New Jersey, Prentice-Hall Inc., 1975.

Guess, D., Sailor, W., Rutherford, G. & Baer, D. M., An experimental analysis of linguistic development: the productive use of the plural morpheme. Journal of Applied Behavior Analysis, 1968, 1, 297-306.

Goetz, E. M. & Baer, D. M., Social control of form diversity and emergence of new forms in children's blockbuilding. Journal of Applied Behavior Analysis, 1973, 6, 209-217.

Guilford, J. P., Creativity, American Psychologist, 1950, 5, 444-454.

Guilford, J. P., Creativity: A quarter century of progress. In I. A. Taylor and J. W. Getzels (Eds.) Perspectives in Creativity, Chicago, Aldine Publishing Co., 1975.

Hersen, M. & Barlow, D. H., Single-case experimental designs: strategies for studying behavior change. New York, Pergamon Press, 1976.

Holman, J., Goetz, E. M. & Baer, D. M., The training of creativity as an operant and an examination of its generalization characteristics.

In B. C. Etzel, J. M. LeBlanc and D. M. Baer (Eds.) New developments in behavioral research: theory and application. In honor of Sydney W. Bijou. Hillsdale, New Jersey, Lawrence Erlbaum Assoc., 1977.

Kendler, H. H., Environmental and cognitive control of behavior. American Psychologist, 1971, 26, 962-973.

Kendler, H. H., Glasman, L. D. & Ward, J. W., Verbal labeling and cue training in reversal-shift behavior. Journal of Experimental Child Psychology, 1972, 13, 195-209.

Maloney, K. B. & Hopkins, B. L., The modification of sentence structure and its relationship to subjective judgments of creativity in writing. Journal of Applied Behavior Analysis, 1973, 6, 425-435.

Parsonson, B. S. & Baer, D. M., Training generalized improvisation of tools by preschool children. Journal of Applied Behavior Analysis, 1978, 11, 363-380.

Skinner, B. F., Science and human behavior, New York, Macmillan, 1953.

Skinner, B. F., The technology of teaching, New York, Appleton-Century-Crofts, 1968.

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

Taylor, I. A., A retrospective view of creativity investigation. In I. A. Taylor and J. W. Getzels (Eds), Perspective in Creativity, Chicago, Aldine Publishing Co., 1975

Torrance, E. P., Curiosity of gifted children and performance on timed and untimed tests of creativity. Gifted Child Quarterly, 1969, 13, 155-156.

Appendix A

Description of experimental materials used as exemplars, potential tools and distractors.

Tool task: Containers
80, 1.27 cm diameter marbles

Baseline Exemplar: Plastic bag with draw string, 22.9 cm x 29.2 cm

Potential Tools:

Item Description	Dimensions
plastic flower pot	10.5 cm dia. x 8.9 cm high
shoe box	33 cm x 16.5 x 10.5 cm
square plastic flower pot	10.5 cm x 10.5 cm x 7.6 cm
glass canister	10.5 cm x 7.6 cm x 7.6 cm
styrofoam cup	8.9 cm dia. x 10.5 cm
cloth carpenter's apron	27.9 cm x 43.2 cm
plastic cycling water bottle	7.0 cm dia. x 17.9 cm
plastic piggy bank	17.9 cm x 10.5 cm
plastic rain hat	55.9 cm x 38.1 cm
coffee bag	454 g
plastic pencil case	21.0 cm x 10.8 cm
beer glass	340 ml
cloth pillow case	83.8 cm x 48.3 cm
8 sided metal electrical connection box	each side 8.9 cm
milk carton	1 litre
milk carton	250 ml
plastic soap dish	13.9 cm x 10.5 cm x 2.5 cm
piece of tin foil	30.5 cm x 45.7 cm
pastry bag	30.5 cm
cylindrical Tupperware container	1.35 L.
leather baseball glove	
cardboard box	10.5 cm x 8.9 cm x 9.5 cm
ladies purse	30.5 cm x 30.5 cm
muffin tray	29.2 cm x 17.8 cm
roller paint tray	35.6 cm x 25.4 cm
metal treasure chest	15.2 cm x 10.5 cm x 10.5 cm
file folder	30.5 cm x 45.7 cm
soap dish	11.4 cm x 9.5 cm x 1.3 cm
man's glove	
plastic sand pail	8.9 cm dia. x 10.5 cm
coffee can	454 g

Item Description	Dimensions
squeeze mustard container	7.6 cm dia. x 7.6 cm
frisbee	22.9 cm dia.
soup bowl	12.7 cm dia. x 5.1 cm
wicker basket	22.9 cm x 17.1 cm x 6.4 cm
long playing record jacket	
triangular Tupperware container	4.4 cm deep x 12.7 cm high x 11.4 cm wide
brown velvet bag	25.4 cm x 15.2 cm
coffee mug	7.6 cm dia. x 8.9 cm
plastic thermos	7.6 cm dia. x 22.9 cm
plastic bag	33 cm x 43.2 cm
glass jar	7.6 cm dia. x 10.5 cm
Big Mac container	
plastic ice cube tray	27.9 cm x 10.8 cm
ceramic ashtray	22.9 cm x 15.2 cm
bathing cap	21.6 cm x 17.9 cm
Man's sock	
foil pie plate	22.9 cm dia.
wool toque	20.3 cm x 25.4 cm
racquetball racquet cover	30.5 cm x 20.3 cm
hand towel	45.7 cm x 61.0 cm
toaster cover	27.9 cm x 12.7 cm x 17.9 cm
brown paper bag	22.9 cm x 6.4 cm x 12.7 cm
plastic folder	40.6 cm x 27.9 cm
brown envelope	22.9 cm x 30.5 cm
chamois	30.5 cm x 45.7 cm
plastic plunger	11.4 cm dia. x 3.8 cm
plastic bag	11.4 cm x 40.6 cm
plastic file/recipe box	8.9 cm x 7.6 cm x 12.7 cm
ladies nite cap	
plastic desert dish	10.2 cm dia. x 3.8 cm
foil sunflower seed package	17.9 cm x 12.7 cm
child's cotton sun hat	
child shoe	size 12
conical party hat	17.9 cm x 12.7 cm dia.
plastic funnel	12.7 cm dia. x 7.6 cm
cooking pot	15.2 cm dia.
deep fry basket	20.3 cm dia.
metal tea pot	2 cup
Volkswagon hub cap	
colander	
honey tub	454 g
ceramic frog	10.2 cm doa. x 10.2 cm

Item Description	Dimensions
thermos top	10.2 cm dia. x 3.2 cm
man's leather mitt	24.1 cm x 12.1 cm
paper plate	17.8 cm dia.
metal can	5.1 cm dia. x 19.0 cm
Cooper #1 wood golf club cover	
bounce ball balloon	27.9 cm long
X-mas wrap	30.5 cm x 30.5 cm
man's cap	18.4 cm dia. x 12.7 cm
Tonka dump truck	33 cm x 15.2 cm

Distractors:

Item Description	Dimensions
kleenex tissue	20.3 cm x 22.9 cm
styrofoam cup	7.0 cm dia. x 8.3 cm
jar lid	1.3 cm x 8.9 cm dia.
man's swim trunks	
plastic flower pot	5.7 cm x 5.7 cm
cardboard box	12.7 cm x 3.8 cm x 3.8 cm
ice cream cone	12.7 cm long
conical party hat	6.4 cm dia. x 9.5 cm
tart cup	7.0 cm dia. x 1.9 cm
envelope	13.3 cm x 11.4 cm
pill bottle	1.9 cm dia. x 5.1 cm
plastic bag	30.5 cm x 12.1 cm
plastic drink coaster	7.0 cm dia. x .64 cm
brown paper bag with hole	22.7 cm x 13.3 cm
mesh purse	33.0 cm x 43.2 cm
paper sheet	12.7 cm x 21.6 cm
change purse	7.6 cm x 10.2 cm
child sock	
leather eye glass pouch	15.2 cm x 7.0 cm
honey tub cover	12.1 cm dia. x .64 cm
36 slide container	
bottle cap	2.54 cm dia. x 1.9 cm
35 mm film canister	4.4 cm dia. x 5.1 cm
cup cake cup	8.9 cm dia.

Tool Task: Hammers
Playskool Cobbler's Bench

Baseline Exemplar: Wooden mallet hammer

Item Description	Dimensions
Potential Tools:	
wood ruler	30.5 cm long
garden shovel	30.5 cm long
wood rolling pin	40.6 cm long
man's cross-country ski boot	size 8 1/2
wood backed chalk eraser	12.7 cm x 5.1 cm
hard cover book	15.2 cm x 22.9 cm
wood meat mallet	30.5 cm long
metal pinking shears	21.6 cm long
wood rod	2.54 cm dia. x 20.3 cm
wrapped bar of soap	133.3 g
plastic ice skate blade guard	30.5 cm long
metal wood-clamp	10.2 cm long
glass candle holder	10.2 cm high
metal wood-plane	16.5 cm x 4.4 cm
metal ice cream scoop	16.5 cm long
flat bastard (file)	31.8 cm long
plastic banana	20.3 cm long
four point deer antler	33.0 cm long
metal bowl	11.4 cm dia.
metal flashlight	17.8 cm long
metal cake pan	22.9 cm x 22.9 cm
can of putty (full)	227 g
empty drink can	284 ml
wooden based, metal bowling trophy	12.7 cm high
plastic bottle opener	11.4 cm long
tree branch	3.8 cm dia. x 25.4 cm
wooden sanding block	11.4 cm x 5.1 cm
empty soft drink bottle	280 ml
metal pliers	15.2 cm long
rubber hockey puck	7.6 cm dia.
box of staples (full)	10.2 cm x 3.8 cm
plastic hairbrush	20.3 cm long
30 mm open ended wrench	25.4 cm long
lead pipe	1.9 cm dia. x 15.2 cm
door knob assembly	15.2 cm x 5.1 cm
"D" size battery	
wooden spoon	30.5 cm long
padlock	6.4 cm x 5.1 cm
hand held drill	30.5 cm long
ping-pong paddle	27.9 cm long
telephone receiver	20.3 cm long
soup ladle	27.9 cm long
adjustable wrench	15.2 cm long
wood chisel	14.0 cm long
metal geometry set box	15.2 cm x 7.6 cm

Item Description	Dimensions
box of 2 inch common nails	454 g
candle	7.6 cm dia. x 15.2 cm
wooden handled paint brush	22.9 cm x 2.0 cm
plasticene ball	7.6 cm dia.
plastic apple	8.9 cm dia.
butter knife	22.9 cm long
glass salt shaker	3.8 cm dia. x 7.6 cm
rock	7.6 cm dia.
screwdriver	20.3 cm long
glass ashtray	14.0 cm x 14.0 cm
croquet ball	7.6 cm dia.
stapler	12.7 cm long
box of playing cards	8.9 cm x 1.9 cm
foil ball	6.4 cm dia.
cork top	6.4 cm dia. x 3.2 cm
glass jar	10.2 cm x 7.6 cm dia.
coffee mug	

Distractors:

Item Description	Dimensions
soft plastic hammer	17.8 cm long
wood block	10.2 cm x 1.3 cm x .64 cm
ball of X-mas wrap	6.4 cm dia.
plastic ruler	15.2 cm long
plastic wrench	15.2 cm long
plastic fork	15.2 cm long
foam rock key-chain	10.2 cm dia.
sponge	16.5 cm x 8.9 cm x 4.4 cm
suede brush	14.0 cm long
cork	3.8 cm x 3.8 cm
garden hose	1.3 cm dia. x 30.5 cm
candle	.64 cm dia. x 17.8 cm
plastic toy bowling pin	15.2 cm long
drawer handle	.30 cm dia. x 11.4 cm
pencil	12.7 cm long
stone	2.5 cm dia.
"AA" size battery	
plastic handled paint brush	18.4 cm x .60 cm
plstic cup	7.6 cm dia. x 6.4 cm
plastic knife	20.3 cm long
tennis ball	

Appendix B

Order of exemplars presented in exemplar training strategies, by child and tool task.

Delia:

Hammers	Containers
1. door knob assembly foam rock	
2. wood ruler cork	
3. plasticene ball garden hose	
4. salt shaker foam rock	
5. butter knife plastic hammer	
6. foil ball plastic wrench	
7. cork top suede brush	
8. hairbrush small candle	
9. ashtray "A" size battery	1. man's mitt styrofoam cup
10. adjustable wrench plastic fork	2. shoe small flower pot
11. book sponge	3. paper sheet mesh purse
12. geometry set box tennis ball	4. bathing cap small soap dish
13. shoe pencil	5. velvet bag 35 mm film canister
14. open ended wrench drawer handle	6. hand towel plastic bag (small)
15. flat bastard ball of X-mas wrap	7. brown envelope small envelope

Hammers

Containers

8. baseball glove
   cup-cake cup

9. paper bag
   paper bag with hole

10. man's glove
    eye glass pouch

11. coffee can
    kleenex tissue

12. quart container
    child sock

13. file folder
    swim trunks

14. pastry bag
    honey tub lid

15. lady's purse
    small cardboard box

16. large styrofoam cup
    change purse

17. Tonka dump truck
    jar lid

18. pencil case
    ice cream cone

19. pillow case
    tart cup

20. narrow plastic bag
    party hat

21. plunger
    paper sheet

Tanya:

1. door knob assembly
   foam rock

2. "D" size battery
   cork

Hammers

3. plasticene ball
   plastic wrench

4. deck of cards
   plastic hammer

5. ashtray
   garden hose

6. drink can
   suede brush

Jason:

Containers

1. egg carton
   soap dish

2. shoe
   small flower pot

1. shoe
   small flower pot

2. paper sheet
   kleenex tissue

3. toaster cover
   35 mm film canister

4. man's sock
   styrofoam cup

5. baseball glove
   plastic bag

6. brown envelope
   small envelope

7. paper bag
   cup-cake cup

8. pastry bag
   honey tub lid

9. treasure chest
   tart cup

10. basket
    eye glass pouch

Hammers

Michelle:

Containers

1. X-mas wrap
   styrofoam cup
2. baseball glove
   small flower pot
3. shoe
   small soap dish
4. toque
   change purse
5. file folder
   small box
6. tea pot
   35 mm film canister
7. colander
   jar lid

Daniela:

1. paper wrap
2. paper plate
3. soap dish
4. long envelope
5. baseball glove
6. toque
7. treasure chest
8. jar
9. colander
10. paper cup
11. glass canister

Alvin:

Hammers

1. "D" size battery
   foam rock

Containers

Hammers

2. butter knife
   cork
3. foil ball
   garden hose
4. screwdriver
   plastic hammer
5. deck of cards
   suede brush
6. candle
   plastic wrench
7. meat mallet
   pencil
8. rolling pin
   plastic fork

Containers

1. shoe
   small flower pot
2. baseball glove
   35 mm film canister
3. brown envelope
   plastic bag
4. man's glove
   styrofoam cup
5. pastry bag
   cup-cake cup
6. file folder
   torn bag
7. pie plate
   change purse
8. large styrofoam cup
   small envelope
9. Tonka dump truck
   small box

Aysa:

Hammers

1. butter knife
2. foil ball
3. cork top
4. pliers
5. meat mallet

Containers

1. shoe
2. X-mas wrap

Hammers

Containers

3. toaster cover
4. paper plate
5. man's sock
6. brown envelope
7. towel
8. frisbee
9. man's glove
10. pie plate
11. thermos
12. basket
13. lady's purse
14. glass canister

Ché:

Hammers

1. deck of cards
2. cork top
3. meat mallet
4. book
5. box of staples

Containers

1. shoe
2. toaster cover
3. X-mas wrap
4. sun hat
5. man's sock
6. hand towel
7. brown envelope
8. paper bag
9. man's glove
10. file folder
11. Big Mac container
12. quart container

Hammers

Containers

13. pastry bag
14. carpenter's apron
15. pastry bag
16. record jacket
17. ashtray
18. pillow case
19. narrow plastic bag
20. bouncing ball balloon

Wendy:

Hammers

Containers

1. soup bowl
2. coffee mug
3. man's mitt
4. shoe
5. X-mas wrap
6. sun hat
7. toaster cover
8. man's sock
9. brown envelope
10. baseball glove
11. paper bag
12. file folder
13. treasure chest
14. quart container

1. garden shovel
2. "D" size battery
3. doorknob assembly
4. butter knife
5. candle
6. meat mallet
7. candle holder
8. can of putty

Hammers

Sylvia:

1. door knob assembly
2. garden shovel
3. wood ruller
4. plasticene ball
5. butter knife
6. foil ball
7. salt shaker
8. cork top
9. screwdriver
10. ashtray
11. paintbrush
12. candle
13. book

Containers

1. shoe
2. toaster cover
3. X-mas wrap
4. man's sock
5. velvet bag
6. hand towel
7. baseball glove
8. man's glove
9. file folder
10. pie container
11. treasure chest
12. pastry bag
13. sunflower seed package