8701324

Devany, Jeanne Marie

STIMULUS EQUIVALENCE AND LANGUAGE DEVELOPMENT IN CHILDREN

The University of North Carolina at Greensboro PH.D. 1985

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

STIMULUS EQUIVALENCE AND LANGUAGE DEVELOPMENT IN CHILDREN

by

Jeanne M. Devany

A Dissertation submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
1985

Approved by

Rosemery O. Nelson
Dissertation Adviser

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Dissertation Adviser Rosemery O. Nelson

Committee Members Aaron J. Brownstein

P. Scott Lawrence

David H. Reilly

9-3-85

Date of Acceptance by Committee

6-20-85

Date of Final Oral Examination

ABSTRACT

DEVANY, JEANNE M., Ph.D. Stimulus equivalence and Language Development in Children. (1985)
Directed by Dr. Rosemery Nelson. Pp. 74.

The stimulus equivalence paradigm offers behavior analysis an approach to the study of semantics. To date, however, no studies of the relationships between language development and stimulus equivalence have been done. Three groups of children, matched for mental age, were studied. One group consisted of normally developing preschoolers, the second consisted of retarded children who used speech spontaneously and appropriately, and the third consisted of retarded children who did not use speech or signs for communication. All children were taught a series of four conditional discriminations and then were tested to determine if classes of equivalent stimuli had formed. All of the language-able children (retarded and normal) formed equivalence classes while none of the language-disabled children did so. Follow-up analyses suggested that the failure to form equivalence classes was due to a failure to demonstrate symmetry in the trained conditional relations, although the possibility that the language-disabled children failed to learn the conditional nature of the training tasks could not be ruled out. The results support the view that the ability to form equivalence classes and language development are related; the nature of the relationship has not been specified through the present research. Additional research directed at clarifying the relationship and teaching the skills prerequisite to the development of equivalence classes is suggested.

ACKNOWLEDGMENTS

The author wishes to thank the staff and students of the Henry Wiseman Kendall Center and the UNC-G Child Care Education Center, especially Joan Moran and Pieter VanIderstine, for their cooperation and support of this project. The author also wishes to thank the members of her doctoral committee for having the patience to read and discuss two completely different doctoral proposals and for their assistance in revising this project. Special thanks are also extended to my advisor, Rosemery O. Nelson, Ph.D., to Rick Cook, who collected all of the reliability data for the project, and to Steven C. Hayes, Ph.D., for his many helpful comments and suggestions.

TABLE OF CONTENTS

Page

APPROVAL PAGE ii

ACKNOWLEDGMENTS iii

LIST OF TABLES. vi

LIST OF FIGURES vii

CHAPTER

I. INTRODUCTION 1

Behavior Analytic View of Meaning. 2
Words as Symbols 4
Stimulus Equivalence 6
Defining Features of Equivalence Classes 7
Stimulus Equivalence and Symbols 8
Equivalence in Infra-humans. 10
Identity Matching and Symmetry 12
Equivalence Classes and Language 12
Statement of Purpose 17

II. METHOD 19

Subjects and Subject Identification. 19
Phases 22
Setting. 25
Procedure. 25
Recording and Reliability. 27

III. RESULTS. 29

Conditional Discrimination Training. 29
Equivalence Test 31
Supplementary Data 34

IV. DISCUSSION 39

Failure to Demonstrate Stimulus Equivalence. 40
Acquisition of the Conditional Relationships 43
Improvement During Testing 44
Generality of the Results. 46
Equivalence and Language 48
Limitations of the Present Study 51
Summary. 52

Page

BIBLIOGRAPHY. 54

APPENDIX A. FIGURES. 60

APPENDIX B. TABLES 73

LIST OF TABLES

Table | Page

1 Subject characteristics 73

2 Reliability data 74

LIST OF FIGURES

Page

Figure

1 The stimulus equivalence paradigm (cf. Sidman, Cresson, & Willson-Morris, 1974). 60

2 The stimulus figures used in this experiment 61

3 The number of trials to mastery criterion during the conditional discrimination training. 62

4 Individual training and testing data for Alex, Allen and Andrew. 63

5 Individual training and testing data for Bobby, Beth and Barb. . 64

6 Individual training and testing data for Claire, Carl and Craig. 65

7 Individual training and testing data for Diane, David and Debbie 66

8 Responding during the test phase 67

9 Testing data 68

10 Suplementary data collected with Craig 69

11 Retraining and testing data collected with Debbie and Randy. . . 70

12 Stimulus set used for training and testing Beth. 71

13 Rachel's "test". 72

CHAPTER I

INTRODUCTION

There is nothing obvious about the nature and function of language.

--E. Lenneberg (1969)

The study of language acquisition focuses extensively on the acquisition of word meaning. These discussions typically focus on the child's increasing ability to "manipulate symbols" (Clark & Clark, 1978) or ability to "map words onto internal concepts" (Nelson, 1974). Words may be said to "stand for" or "refer" to objects, events, and the relationships between and among them (Gentner, 1978; Premack, 1983).

Although in traditional views of language much is made of the the symbolic nature of words, we still do not understand why or how words come to function as symbols. In fact, in one of the books mentioned above (Clark & Clark, 1977), which by two eminent psycholinguists, no index reference for "symbol" or "symbols" was found. It is as if the symbolic nature of words is a "given" and the literature then traces the use of words/symbols through the developmental period. This lack of attention is probably related to the views of language held by non-behavioral theorists. A view that emphasizes the nativistic aspects of language abilities would seek to document these special abilities but not to demonstrate their ontogenetic origins. Similarly, research done within a traditional developmental perspective would seek to identify the stages in the growth of language ability but not to identify the conditions that give rise to them. Yet it is apparent that phylogenetic endowment alone is insufficient to guarantee the emergence of language. Children raised in isolation, such as the child known as Genie (Curtiss, 1981), do not develop speech or other language systems spontaneously. A complete understanding of the development of language requires an understanding of the experiences that lead to the use of words as symbols and other aspects of language development.

An alternative point of view may be found within the behavior analytic framework. It is characteristic of this perspective that events are understood functionally and an identification of the conditions that give rise to the development of symbolic behavior might be expected to emerge most readily from this paradigm. In this introduction, the behavior analytic account of meaning and the limitations of this account are described. A new concept, stimulus equivalence, is described. This concept appears to account for or capture some of the sense of "word meaning". Finally, the rationale for a project investigating the relationship between stimulus equivalence and language is described.

Behavior Analytic View of Meaning

Within a Skinnerian framework, the meaning of a word is found in the practices of the language community within which the word is used. That is, "meaning" is not found in the topography or structure of a word; it is not a property of the word per se. The meaning of the arbitrary vocal stimulus "bread" is found through observation of the circumstances that occasion its emission and the circumstances that do not. When "bread" is a tact, the common physical properties found across situations in which bread is emitted define "bread" (Skinner, 1957). A referent, then, is an object/event/relation that is conditionally related to an arbitrary stimulus (the symbol): If bread, then "bread". Or, to put it another way, in the presence of a given stimulus, the probability of a given response is increased.

In this connection, Skinner (1974) uses the example of a rat trained to press a bar. In the presence of a flashing light, bar-pressing is consequated with water and in the presence of a steady light, bar-pressing is consequated with food. Skinner argues that it is possible to say that the flashing light "means" water and the steady light "means" food. Further, he argues that (for the speaker) the relationship between the environment and what is uttered is the same as the relationship between the lights and the bar-pressing. (He allows that the meaning of an utterance is the conditional relationship in combination with all other relevant contingencies but the essential point is the same.)

Yet, the conditional relationship does not appear to be a completely satisfactory description of the word-referent relationship. The relationship between seeing a mosquito on one's arm and swatting the arm is a conditional relationship: If mosquito, then swat. Few would argue, though, that this is a linguistic/symbolic relationship. Thus, there seems to be a difference between the symbolic relationship and a conditional discriminative relationship. Most students of language, for example, would agree that the word "bread" stands for or refers to an object (e.g., Ristau & Robbins, 1982). It is nonsensical to say, however, that the act of swatting "stands for" the mosquito.

Behaviorists have generally denied that verbal stimuli differ from other discriminative stimuli. For example, a conditional relation exists between the command "Come here" and a dog's approach to its owner. The action of the dog is similar to the behavior of a human being when told "come here". Superficially there appears to be nothing special about the control exerted by the verbal stimuli over the human as compared with the control exerted over the dog. The argument is made, then, that either all discriminative stimuli are symbols or that the entire issue of word meaning is cast around an unnecessary concept (symbols).

Words as Symbols

The word "symbol" comes from root words meaning "together or alike" and "to throw". Symbols, then, are stimuli that are "thrown together" with other stimuli. As mentioned earlier, when we talk about symbols, we imply that a one stimulus "stands for" another and vice versa. In other words, the relationship appears to be bi-directional.

This issue has been extensively discussed in the controversy surrounding language acquisition in apes. Washoe, the chimpanzee trained to use American Sign Language by the Gardners (Gardner & Gardner, 1969), has learned a number of labels for a variety of objects and to indicate a variety of wants. Yet considerable debate exists as to whether Washoe and other language trained pongids (e.g., Fouts, Chown, & Goodin, 1976; Patterson, 1978) use these symbols in the symbolic/referential manner ascribed to human children or whether these are only unidirectional conditional relationships (Bronowski & Bellugi, 1970; Chomsky, 1969; Lieber, 1982; Muncer & Ettinger, 1984; Sebeok, 1980). In other words, "bread" signed by a chimpanzee is thought by some workers in the field to be merely a response reliably made in the presence of bread (or, perhaps, in the presence of hunger)-- a conditional relationship rather than a "true" symbolic relationship.

Perhaps another example will illustrate the point. In a recent study (Pepperberg, 1983), a parrot was taught to utter color names (e.g., red or green) when presented with an object and the question "What color is it?" and to utter shape names when presented with an object and the question "What shape is it?". The behavior of the parrot, although conditionally related to the questions/objects, is not "language" in the commonly accepted use of the term. Most people have no difficulty reaching this conclusion in spite of that fact that no universally accepted definition of language is available. When asked, most people would say that the parrot does not "understand" what it is saying or that the words are not used as "symbols." (In fact, we perjoratively characterise utterances in which the speaker seems to speak without understanding as "parroting"). It is often difficult to point out precisely why the performances of the parrot and the chimp do not qualify as true linguistic performances. Two of the major contributors to this problem are a lack of general agreement on what constitutes "true" symbolic behavior and by a failure of researchers in the area of child language to examine adequately symbol use in the developing child.

Words/symbols used by humans appear to be much more flexible than the words used by chimps or parrots. For example, in a Quebecois human child the English written word "bread" or the French written word "pain", their spoken counterparts, pictures of bread, and loaves of bread all enter into a rich network of relations in which each may (in a sense) stand for the others. If this "interrelatedness" is part of the defining character of symbols, it is not clear that chimpanzees or other animals use sign labels (or other communication devices) as true symbols, as it is argued that normal children do.

Stimulus Equivalence

There is one behavioral process that appears to relate quite closely to the issue of symbolic activity. Sidman and his colleagues have suggested that word meaning may be analyzed in terms of a concept known as stimulus equivalence (Mackay & Sidman, 1984; Sidman & Tailby, 1982). When humans are taught a series of related conditional discriminations, it often happens that the component stimuli of the discriminations become related to each other in new ways, not explicitly taught in training. For example, if a child is taught to match A to B and then A to C, it is likely that the child will, without additional training, be able to match B to A and C to A and, perhaps most importantly, to match B to C. That is, if a child is taught to match the spoken word "bread" to a loaf of bread and later is taught to match the printed word BREAD to the loaf of bread, upon testing we will find that the child is able to match the loaf of bread to the spoken word "bread" and to match the loaf of bread to BREAD as well as to match "bread" to the printed word BREAD without additional training.

In the stimulus equivalence paradigm, a symbol and its referents form a class of functionally substitutible elements. The relationship between a symbol and its referent is not a unidirectional conditional relation (although the members of the class are conditionally related to each other); the relation is functionally reversible. In addition, untrained relationships between members of an equivalence class appear. In one experimental demonstration of stimulus equivalence (diagrammed in Figure 1; this and all subsequent Figures are in Appendix A), two severely retarded institutionalized adolescents were taught to match printed words (e.g., CAT) to dictated words ("cat") and pictures to dictated words. Following this training, the subjects were able to match printed names (CAT) with the appropriate pictures and vice versa although they had not been specifically taught to do so (Sidman, Cresson, & Willson-Morris, 1974).

Defining Features of Equivalence Classes

An equivalence class is said to exist if the stimuli in the class show the three defining relations of reflexivity, symmetry, and transitivity (Sidman & Tailby, 1982). In matching-to-sample procedures, reflexivity is generalized identity matching-- matching a novel stimulus to itself under conditions of no reinforcement. That is, if an organism is presented with a piece of bread, it will select another piece of bread from an array of items even in the absence of programmed reinforcement for that choice.

Symmetry refers to the functional reversibility of the conditional relation. If A, then B; and if B, then A. In the presence of the printed word BREAD, a loaf of bread is selected; and in the presence of a loaf of bread, the printed word BREAD is selected. Sidman and his colleagues (Sidman, Rauzin, Lazar, Cunningham, Tailby, & Carrigan, 1982) have asserted that this reversibility must be demonstrated in the absence of direct reinforcement to be considered symmetry.

To demonstrate transitivity, which is the emergence of untrained relations (the dotted lines in Figure 1), at least three stimuli are required. If after the relations "if A, then B" and "if B, then C" have been taught, the relation "if A, then C" emerges without additional training, transitivity has been demonstrated. In other words, if the child has been taught to select bread upon hearing "bread" and to select the printed word BREAD upon seeing bread, when the child selects BREAD upon hearing "bread", without any additional teaching, then transitivity has been demonstrated.

Stimulus Equivalence and Symbols

The relationships among the members of an equivalence class appear to approximate what psycholinguists and others mean when they say that a word represents or "stands for" its referent in a way that a conditionally related response does not. That is, the symmetrical relationship between "bread" and bread permits the listener to act (in some contexts) with respect to "bread" as if it were bread. For example, if one wished to learn to make bread, it would be possible to do so by observing a baker in action. Alternatively, it may suffice to refer to a cookbook under the heading BREAD. Further, behavior learned in one condition may be exhibited appropriately under novel circumstances if the subject has a particular history with respect to the behavior and context. For example, if round loaves of French bread, salt sticks, Wonder bread, and pita bread have all previously been selected in response to the utterance "bread" , the round loaves, the Wonder bread (and so on) will be matched appropriately, even if there is no history of direct reinforcement for doing so.

If equivalence classes have to do with symbolic activity, and if symbolic activity is characteristic of language and yet different from mere discriminative control, we would expect that it would be easy to demonstrate stimulus equivalence with humans and difficult to demonstrate stimulus equivalence with nonhumans. This is exactly what has been found in the research done to date. The formation of classes of equivalent stimuli in humans has been examined in a number of studies.

The formation of equivalence classes has been demonstrated using matching-to-sample techniques with retarded adolescents (Sidman, 1971; Sidman et al., 1974), young retarded children (Spradlin, Cotter, & Baxley, 1973; Spradlin & Dixon, 1976), and young normal children (Sidman & Tailby, 1982), using both auditory and visual stimuli (Dixon, 1976), visual stimuli alone (Wetherby, Karlan, & Spradlin, 1983), and auditory stimuli alone (Karlan, 1976). "Real" words (symbols) and pictures (referents) have been used (Sidman, 1971) as well as arbitrary visual referents and nonsense words (Dixon, 1976). Some studies have employed numerical symbols and number-groups (Gast, vanBiervlet, & Spradlin, 1979) while others have used manual signs and objects (vanBiervlet, 1977). All of these studies followed the general outline of the Sidman et al. project (described earlier and depicted in Figure 1), and the Lazar et al. project, although procedural details varied from study to study.

Equivalence in Infra-humans

While conditional relationships have been demonstrated in a large variety of lower animals, including dolphins (e.g., Herman & Thompson, 1982), rats (e.g., Lashley, 1938), pigeons (e.g., Edwards, Jagielo, & Zentall, 1983), monkeys (e.g., Nissen, 1951), and mynah birds (Turney, 1982), these do not result in stimulus equivalence as they do in most humans. To date, there has been no success in unequivocally demonstrating the formation of an equivalence class in any infra-human, including the higher primates, although efforts have been made (Sidman et al., 1982). Further, researchers have had considerable difficulty demonstrating transitive transfer even under conditions of direct reinforcement in chimpanzees (Fouts et al., 1976) and other animals (Kendall, 1983).

In the Fouts study, a young male chimpanzee (Ally) was taught to produce American Sign Language sign labels in response to vocal English labels. He had previously learned to identify, by pointing or touching, these objects when given the appropriate vocal command. For example, he had learned to select his pillow from an array of his toys when presented with the vocal statement "pillow". Later, in the absence of pillows or other toys, he was taught to produce the sign label pillow in response to the vocal statement "pillow". Once he had mastered this, a "test" phase was administered in which the physical object (e.g., the pillow) was presented and Ally was asked in sign language to identify it. Praise and tickling were provided for correct responding so this cannot be considered a true test of transitivity. Even with the rewards, however, as many as 14 hour-long test sessions were required before Ally reliably produced the correct sign label when presented with an object. Thus, even with reinforcement for correct responding, it proved difficult to demonstrate transfer in an infrahuman primate who was apparently motivated and had already mastered an 80 word vocabulary.

Pigeons perform even more poorly on transfer tests. In Kendall's 1983 study, pigeons each were taught two conditional discriminations. For example, a bird was taught to peck either the left or the right front wall key in the experimental chamber, depending on which of two signal lights (white or amber) was lit. Both of the front wall keys were illuminated red. In the second discrimination, keys on the side wall of the experimental chamber were used; and pecks on the left key or the right key (both illuminated green) were reinforced depending on which of the two signal lights was lit. In the test phase, one of the front wall keys was illuminated. After the bird pecked the front wall key, the side wall keys were illuminated. To produce grain in the feeder, the bird was required to peck the side wall key that "went along with" the illuminated front wall key based on the previous conditional discrimination training. Neither of the two birds trained in this manner performed above chance on the transfer test. The procedure was modified for three other birds with equally poor results. This was true even though in some cases the transfer testing was carried out over many sessions.

Identity Matching and Symmetry

In fact, the demonstration of generalized identity matching in monkeys (D'Amato, Salmon, & Columbo, 1985) and other animals (Carter & Werner, 1978) remains controversial. In the D'Amato et al. study, monkeys did develop a generalized matching concept after being taught visual identity matching with a small number of exemplars. However, the range of situations in which the monkeys spontaneously applied the matching concept was quite restricted relative to young children, even though the novel stimuli were in the same modality as that in which the original training occurred. In addition, although efforts have been made to find symmetry in conditional discriminations taught to monkeys and baboons (Sidman et al., 1982), these efforts have been completely unsuccessful to date. The failure to demonstrate equivalence classes in lower animals, the difficulty in demonstrating transfer even with reinforcement, and the failure to demonstrate symmetrical relations, as well as the difficulty in demonstrating generalized identity matching, all make more plausible the hypothesis that the ability to form stimulus classes is related to language development.

Equivalence Classes and Language

If equivalence classes have to do with language, we would expect the control that members of equivalence classes exert over behavior to parallel the control language exerts over behavior. The available evidence indicates that this is what occurs. Once a set of stimuli has become an equivalence class, each member of the class becomes substitutible within certain contexts. For example, if visual stimuli A, B, and C form an equivalence class and A is taught as a response to the auditory stimulus "zug", then B and C will also be chosen in response to "zug" (Anderson & Spradlin, 1980; Dixon & Spradin, 1976; Spradlin & Dixon, 1976). Or, if trains, dolls, and puzzles form an equivalence class, once the child is taught to pick up her doll in response to "toy", trains and puzzles will also be selected in response to "toy".

In addition, after an equivalence class (ABC) is formed, if one member (A) becomes discriminative for a response (Z), then B and C will also become discriminative for the emission of Z (Hayes, Brownstein, Devany, Kohlenberg, & shelby, 1985; see also Lazar, 1977). For example, if Pop-tarts, potato chips, and raisins form an equivalence class and the child is taught to say "snack" in the presence of potato chips, the child will also say "snack" in the presence of raisins and Pop-tarts, even though she or he had never been explicitly taught to do so.

These findings appear related to the work done on "semantic generalization". For example, if a word is paired with an unconditioned stimulus, such as a puff of air, in a classical conditioning paradigm, words related to the conditioned stimulus will elicit the conditioned response, but words that sound similar to the conditioned stimulus but "mean something different" will not. If the conditioned stimulus is "doctor", words such as "physician" will elicit the conditioned response but similar sounding words (e.g., diktor) will not (Shvartz, 1960, described in Slobin, 1971). Interestingly, very young children respond to words of similar sound rather than words of similar meaning in such experiments. The explanation offered for semantic generalization has been that "... clusters of words have become associated on semantic grounds, (and thus,) responses established to one part of the cluster can be elicited by other parts of the cluster" (Slobin, 1971, p. 84). The problem is that a satisfactory definition of what it means to "become associated on semantic grounds" has not been been developed. It seems plausible to suggest that the "association" mentioned here is participation in equivalence classes.

We also find, in the equivalence class literature, data that parallel observations made on the acquisition of language. For example, much of our knowledge of words and their meanings undoubtably comes from observational learning in addition to direct tuition. If stimulus equivalence is closely related to words/symbols, we would expect to find that a stimulus could become a member of an equivalence class through observational learning. This has been found to be true (MacDonald, 1983). When retarded adults were taught one visual conditional discrimination (e.g., A-B) directly and observed a classmate learning a related conditional discrimination (e.g., A-C), when tested they were able to correctly match B and C and to show symmetry in both the directly trained and the observed conditional relationships.

The stimulus equivalence paradigm may also help us to analyze syntactical behavior. In one study (Lazar, 1977), normal adults were taught to point sequentially (first one and then the other) to each member of several pairs of visual stimuli. That is, they were taught to touch A1 and then A2 and B1 and then B2, regardless of their position (left or right) on the stimulus display board. Upon testing, the subjects touched the stimuli in the proper sequence when A2 was paired with B1 and when A1 was paired with B2. The subjects were then taught to select a new stimulus, E1, in the presence of A1 in a matching-to-sample task. They were also taught to select E2 in the presence of A2. When tested, they correctly matched E1 to B1 and E2 to B2 without direct training. And, in a final test, when E1 and E2 were presented in the sequencing task, the subjects touched the stimuli in the order predictable on the basis of their membership in equivalence classes. That is, the subjects touched the stimuli in the correct order without training.

We might think of this as comparable to a child who is taught to say "blue ball" in the presence of the appropriate object and who later utters "yellow ball" in the appropriate context. While in the first instance, the utterance can be traced to the child's reinforcement history without difficulty, it has been difficult for a functional analysis to account for the emission of "yellow ball", as a direct reinforcement history is lacking. This behavior could be understood if the grammatical elements ("yellow" and "ball") are viewed as members of syntactical classes (Jenkins, 1965). That is, words that regularly occupy the same position in a sequence may become members of an equivalence class. When additional members of the class are trained ("yellow"), the new class member may be used in the same way as are the other members of the class. This would allow the child appropriately to combine adjectives and nouns learned under separate conditions and, thus, to produce novel utterances.

From the available evidence, the stimulus equivalence paradigm appears to provide a basis for the analysis of word-referent relations and novel verbal behavior. Although most research in the area of stimulus equivalence has not addressed the specific issues of semantics and language development, it seems likely that the ability to form equivalence classes is intimately related to language use.

However, no experiments relating language use and equivalence classes have yet been done. One approach to the problem could be to compare the performances of normally developing children with the performances of language-impaired children on an equivalence test. If the normally developing children were able to form equivalence classes and the language deficient children were not, the data would provide further support for the view that the ability to form classes of equivalent stimuli is related to language. Although this is a correlational approach and cannot identify the nature of the relationship between stimulus equivalence and language, the identification of differences in performance between groups would support the utility of continued investigation. If, for example, the ability to form equivalence classes is essential to the use of words in a referential and generative manner, it is possible that by teaching language-deficient children the defining features of equivalence classes (generalized identity matching, symmetry and transitivity), improvements in the appropriate and spontaneous use of words could be achieved. This goal is particularly important as the failure to generalize the gains made in language training programs (Harris, 1975) and other behavioral training programs (Stokes & Baer, 1977) continues to be one of the major challenges to the successful habilitation of the developmentally disabled.

Statement of Purpose

This project addressed the question: Is there a difference between normal and language deficient children in performance on tests of equivalence? Because lack of language is often associated with mental retardation, three groups of children were employed: normally developing children, retarded children demonstrating some expressive speech, and retarded language-deficient children. These groups were used to control for the effects of retardation per se on performance. That is, the inclusion of a group of retarded children who had language skills was an effort to insure that differences in performance among the children were due to the absence of language and not to the generalized effect of some organic insult, chromosomal damage, or the relatively deprived social and environmental contexts within which many such children live. The three groups were matched on mental age. While this is not an infallible method of controlling for differences between children, it is a widely accepted technique for insuring similarity in intellectual and adaptive repertoires (Achenbach, 1969, 1970). All of the children were taught a series of conditional discriminations and were subsequently tested to determine if an equivalence class had formed.

The hypothesis that stimulus equivalence and language are closely related would not be supported if the normal children performed well on the equivalence task and both groups of retarded children did not or if all three groups of children performed equally well. Perhaps it would not be completely accurate in the former case to say that the hypothesis was not supported; the ability to form equivalence classes might not appear until after a certain level of linguistic competence has been attained. The assumption here, however, is that equivalence classes are essential to the development of words-as-symbols, and this position would be weakened by failure of the retarded children with some language skills to perform well in the equivalence test. If all three groups performed equally well, it might be reasonable to argue that the ability to form equivalence classes and language abilities are distinct.

It was expected that the normal children would master the conditional discriminations more quickly than the retarded children and that the normal and retarded children with some language skills would perform better than chance on the stimulus equivalence test. It was not expected that the children in the retarded language-deficient group would perform better than chance on the stimulus equivalence test.

CHAPTER II

METHOD

Subjects and Subject Identification

Twelve children, four in each group, served as subjects. The first group consisted of normally developing preschoolers recruited from the UNC-G Child Care Education Center. The second and third groups were composed of retarded children. All of these children were enrolled in educational programs at the Henry Wiseman Kendall Center in Greensboro.

All of the normally developing preschoolers had speech skills that were generally consistent with their chronological ages. No formal assessment of their speech and language skills was done; in the training and testing sessions, however, no abnormalities of speech or language were noted. In addition, no abnormalities were noted by the classroom teacher or were observed during in-class observation by the experimenter.

Half of the retarded children used speech for communication outside of language training sessions. Some of these children had articulation problems, however, (Carl and Allen, names changed to protect confidentiality) and their speech was at times difficult to understand. All of these children spoke in complete, albeit brief, sentences when prompted and often spontaneously asked for desired items or commented on events in the classroom.

The other half of the retarded children lacked functional speech or language skills. None of these children used words, signs, or picture boards consistently to communicate. Two of the children were echolalic, repeating words or phrases without comprehension (Debbie and Andrew), and two of them uttered vowel sounds (Craig and Barb).

The retarded children were classed into the two groups, the retarded/language group and the retarded/no language group, on the basis of converging opinions from three different observers. The speech pathologist at Kendall Center categorized all of the children in the two preschool classes as possessing functional speech or sign skills (used for communication, even if poorly articulated) or as lacking functional speech or sign skills. This was done before the study began. She was not told the purpose of the study other than that it was an investigation of the ease of concept learning in children of varying levels of language and cognitive skills. The experimenter observed all of the children in their Kendall classrooms for a minimum of one and a half hours prior to the onset of the project. Any children classed as language-able by the speech pathologist who was not observed by the experimenter using speech (or signs) appropriately without prompting in the classroom would have been eliminated from the study. Any child classified as language-disabled who was observed by the experimenter using speech or signs appropriately during the classroom observation period or during the experimental sessions would have been removed from the language-disabled group and would have participated in the supplementary analyses that were conducted. In practice, none of these steps was necessary. Finally, the reliability observer used throughout the experiment was an advanced graduate student with many years of experience working with retarded and language-impaired children. He observed each of the children at least once (for approximately an hour) during an experimental session. After the experimental session was completed, he was asked to categorize the child on the basis of his or her behavior in the experimental session. (This was done once for each child). If his assessment in any case differed from the assessment made by the experimenter, then the child would have been replaced with another child. In all cases, however, the opinion of the reliability observer coincided with the opinions of the experimenter and speech pathologist.

The children in the three groups were matched for mental age on the basis of appropriate individual intelligence tests, such as the Stanford-Binet. The retarded children were all assessed by personnel at Kendall Center. The normal preschoolers were assessed by the experimenter.

Parental permission was obtained prior to the children's participation in the project. Parents were reminded that they were free to withdraw the child at any time. Once the child's participation was complete, the parents were sent a letter informing them of the study's results and describing their child's behavior.

For the purposes of this project, the children's names were changed to protect confidentiality. Each child was given a name beginning with the same letter as the names of his or her mental-age-matched controls. The children and their mental and chronological ages are presented in Table 1. (Table 1 and all subsequent tables are located in Appendix B.)

The independent variable was the subjects' classification into one of the three groups.

Phases

Training phase. In the training phase, the children were taught four conditional discriminations: If A, then B; if D, then C; if A, then E; if, D, then F. For example, if a clown picture had been presented as the sample (A), the child would have been praised for touching the picture of the balloon (B) and ignored for touching the picture of the gerbil (C). However, if the picture of the woodchuck (D) had been presented as the sample, the child would have been praised for touching the picture of the gerbil and ignored for touching the picture of the balloon. (The stimuli in this project were not related in such obvious ways. This example is for illustration only.) The tasks consisted of matching nonsense (made-up) animal figures using a matching-to-sample format. On all trials, the sample (either A or D) was presented with two comparisons. That is, on each trial, the A or D stimulus was presented with either B or C as comparisons or E and F as comparisons.

The stimulus sets were made by randomly selecting from a pool of items the stimulus items used with a particular child. The selections were made by numbering the stimuli and placing the numbers in a hat. The first stimulus selected became A, the second B, and so on. Once all the selections for one child had been made, the numbers were replaced, and the items for another child were selected. After the stimulus set had been selected, the individual stimulus presentation sheets were made by photocopying the original stimulus items and cutting-and-pasting them onto blank sheets. All of the stimulus figures were colored with watercolor magic markers. Six colors were used: red, brown, green, purple, yellow, and orange. Each item was colored a different color. Color assignment was made by putting the color names into a hat, and mixing and drawing out the names one at a time. The first color drawn was used to color A, the second B, and so on.

The three stimuli (one sample and two comparisons) were presented on 8 1/2 x 11" sheets of white paper. When the long side of the sheet was placed horizontally, the sample stimulus was at the top center of the page, while the two comparisons were in the bottom half of the page, each approximately 2" from the edge of the sheet. The two comparisons were approximately 3" from each other. The left-right order of presentation of the comparison stimuli was nonsystematically varied across trials to prevent the child from responding correctly based on position cues alone. Each sheet contained the stimuli for one trial. When the trial was completed, the sheet was removed, and a new sheet with the stimuli for a new trial was presented. (A diagram presenting all of the stimulus figures used in the project is presented in Figure 2. A diagram of one training and testing set used in the project is presented in Figure 12.)

Each of the children was taught individually. The training sequence used for each of the children was identical. First, the child was taught to select B in the presence of A (A-B). Then, the child was taught to select C in the presence of D (D-C). Then, these two tasks were mixed; the stimulus cards from both sets were mixed together and presented in a random order. Once this task was mastered, training on A-E was begun. Then D-F was taught. Once D-F was mastered, A-E and D-F were mixed and presented to the child. When the child reached the mastery criterion on this mixed task, the stimulus items from all four conditional discriminations were mixed and presented. Once the child reached the mastery criterion on this final task, the test trials were presented.

The mastery criterion used throughout the training was nine out of ten consecutive responses correct. These responses had to be unassisted responses. There was one main dependent variable in the training phase. This was the number of trials required by each child to complete the training. The number of prompts needed to complete the conditional discrimination training was also recorded.

Test phase. The materials used in the test phase were identical to those used in the conditional discrimination training portion of the study. However, in the test phase, the sample stimuli were stimuli that had previously been comparisons during the conditional discrimination training. In ten trials of the test phase, the sample stimulus was B and the comparison stimuli were E (correct) and F (incorrect). In ten trials, the sample was E and the comparisons were B (correct) and C (incorrect). In ten trials, the sample was C and the comparisons were E (incorrect) and F (correct). In ten trials, the sample was F and the comparisons were B (incorrect) and C (correct).

There was one dependent variable in the test phase. This was the number of correct responses during the forty trial test phase. For the purposes of this project, a correct response during the test was defined as one that would be expected if an equivalence class had been formed.

Setting

The sessions were held in an unused classroom at Kendall Center and in an office at the Child Care Education Center. The classroom was spacious and contained play equipment, sleeping mats, a desk, and child-sized tables and chairs. The sessions were held at a small table which was clear of extraneous objects. The child always sat facing the table which was pushed up against the wall. This insured that the child sat facing the wall while remaining close to the experimental materials, which were placed on the table in front of the child or were held up off the table slightly (to permit easier viewing). The office at the Child Care Education Center was small and contained a large desk, chairs, a filing cabinet, and a bookcase. The children and the experimenter sat on a rug on the office floor. The stimulus sheets were presented by placing them on the floor directly in front of the child or by holding them slightly above the floor.

Procedure

At the beginning of the session, the experimenter greeted the child and spent several moments conversing. Even if the child was nonverbal, the first portion of the session was spent chatting in order to set a relaxed and pleasurable tone. In the initial sessions, the experimenter started off by saying, "I have some things I would like you to help me with. Let's see if you can help me. I also have some things to play with and we will play with those as well." (For the nonverbal children, this was simplified to "Let's do some work".) The first task (A-B) was then presented. Training continued until the child responded correctly on nine out of ten consecutive trials without assistance.

The normal children were rewarded for correct responses with praise and blowing soap bubbles as well as occasional singing and playing with a talking stuffed monkey. Correct responding by the retarded children was consequated by brief access to tiny flashlights, soap bubbles, balloons, juice, and cheese crackers. That is, for all of the children, a correct response led to five or ten seconds of play with the soap bubbles, a few moments of play with a blown-up balloon, the delivery of a small piece of cheese cracker or a sip of juice, and so forth. When necessary, physical prompting (guiding the child's hand to the correct choice) and visual prompting (pointing to the correct choice) were used as teaching aids. When used, they were faded as quickly as possible. Initially, all correct responses led to the delivery of one of the consequences. At the end of the training period, during the time that all four tasks A-B, D-C, A-E, and D-F were mixed and presented, the schedule was gradually thinned until a programmed consequence was delivered only after three or four correct responses.

When a child became irritable or responding became erratic (even with prompting), the session was discontinued. Every effort was made to return the children to their classrooms in good spirits. If a child did not wish to accompany the experimenter, this wish was respected, and the experimenter returned another day.

In the test phase, forty trials were presented. The composition of these trials has already been described. The four trial types were randomly intermixed. After every third or fourth response (correct or incorrect), the child was praised for cooperation, good sitting (etc.)

or one of the other programmed consequences was delivered. If the child asked for explicit feedback about a response, the experimenter said, "In this part of the game, I must be very quiet. I think you are doing a good job of working on this."

Recording and Reliability

Responses were scored as correct, incorrect, or no response. The experimenter was the primary data collector. A correct response was defined as touching the correct comparison stimulus while refraining from touching the incorrect comparison or the sample stimulus. (Because of the poor performances of the retarded/no language children, they were required to touch the sample prior to touching the comparison. In these cases, a response was scored as incorrect if the child touched the sample more then once or if the child touched the sample and then failed to touch the correct comparison.) An incorrect response was defined as touching the incorrect comparison, touching the sample, touching both the correct comparison and the sample or incorrect comparison, or touching another part of the stimulus sheet. No response was defined as any other response.

Reliability data were collected in approximately twenty percent of the sessions, distributed across children. These data were collected by a trained graduate student who was familiar with the general nature of the project but was unfamiliar with the specific hypotheses. (That is, he knew many of the project details but was unable to describe the experimental hypotheses). The rater sat in a position from which he could not observe the experimenter's data sheet. During sessions in which the observer was present, the experimenter paused briefly after each trial before delivering the consequence. This allowed the observer to record the data without knowledge of the experimenter's scoring.

Reliability was calculated on a trial by trial basis using the formula Agreements/Agreements+ Disagreements times 100. An agreement was scored if both of the two observers recorded a response as correct or as incorrect or as no response. For the purposes of reliability, a prompted trial was considered correct. Agreement per session ranged from 88% to 100%. Session by session reliability data are presented in Table 2.

CHAPTER III

RESULTS

As predicted, the language-able children (those in the normal and the retarded/language groups) performed better than the language-deficient children in all areas. The children's performances on the conditional discrimination tasks are described first. Then the performances on the equivalence test are described. Finally, supplementary data analyses are described.

Conditional Discrimination Training

A one-way analysis of variance indicated that there was a significant difference among the three groups in the number of trials needed to complete the conditional discrimination training, $F\ (2,\ 9) = 6.34$, $p < .019$. A Newman-Keuls post-hoc analysis (Ferguson, 1976) found no significant difference between the normal and the retarded/language group (Figure 3). On average the normal children required 165 trials to complete the seven stages of the conditional discrimination training (A-B, D-C, mix, A-E, D-F, mix, and the final mix). The mean for the retarded/language group was 247.75 trials. The retarded/no language group required significantly more trials ($X = 476.75$) than the normal group ($p < .05$) and the retarded/language group ($p < .05$).

An inspection of the individual data confirms that the children in the retarded/language and normal groups consistently required fewer trials to complete the conditional discrimination training than did children in the retarded/no language condition. The individual data are presented in Figures 4, 5, 6, and 7. In each figure, the data for one normal child, one retarded/language child, and one retarded/no language child are presented. The children in each figure are all matched for mental age. The data presented in the first graph (top) in each figure represents the data for the normal child. The second graph (middle) represents the data from the retarded/language child while the third graph (bottom) represents the data for the retarded/no language child. The data are graphed as the percentages of unprompted correct responses in blocks of ten trials. In other words, the trials on which the child made a correct response without assistance were divided by the number of trials administered in that block (always ten) in order to obtain the percentage. The tasks (A-B, D-C, etc.) are indicated along the abscissa.

Scanning across the figures, consistent differences in the performances of the children in the different groups are apparent. That is, the acquisition curves for each of the seven training tasks differ considerably across children, although within each group there is much less variability. Another way of describing the data is in terms of the total number of trials administered before the children reached the criterion on the final training task. In the normal group, Alex required 95 trials to complete the training and Bobby required 107 trials. Claire required 185 trials while Diane required 273 trials. In the retarded/language group, Allen required 277 trials, Beth required 223 trials, Carl required 227 trials, and David required 264 trials. In the retarded/no language group, Andrew required 507 trials, Barb required 280 trials, Craig required 370 trials, and Debbie required 750 trials.

There was also a significant difference among the three groups in the number of prompts used in the conditional discrimination training, $F(2, 9)=5.42$, $p<.029$. Newman-Keuls post-hoc analyses again showed no significant difference between the normal and the retarded/language groups. The mean number of prompts (visual and manual) used in teaching the normal group was 28.5. The mean number of prompts used in teaching the retarded/language group was 40. The retarded/no language group was significantly different from the normal ($p<.05$) and the retarded/language ($p<.05$) groups. The mean number used in teaching the retarded/no language group was 184.25.

<u>Equivalence Test</u>

The data confirmed the prediction that the language-able children would perform significantly better than the language-deficient children on the stimulus equivalence test, $F(2, 9)= 18.51$, $p<.0006$. A Newman-Keuls analysis revealed no significant difference between the normal and the retarded/language group. The average correct responding in the test phase in the normal group was 84.5%. That is, the children responded correctly on average on 84.5% of the total test trials administered. In the retarded/language group, the average percent correct responding was 78.25%. Significant differences were found between the retarded/no language and the normal ($p<.01$) and the retarded/language ($p<.01$) groups. In the retarded/no language group the average correct responding was 44.5%.

The individual data are presented in Figure 8. Each graph represents the data from one child. The data are graphed as the percentages of correct responding in blocks of ten trials. The columns within each individual graph represent the number of "no responses" made by the child during each block of ten trials. Since some "no responses" did occur, the percentages of correct responding were calculated as (the number of correct responses) divided by (the total number of responses in that block). Each row of graphs represents the data from one group. Each column of graphs represents the data for those children matched for mental age (across groups).

The normal and retarded/language children (the top and middle rows) consistently showed an improvement in the number of correct responses made over the course of the forty trial test. This improvement was not seen in the retarded/no language group. The performances of the children in the retarded/no language group remained at chance level (50%) throughout the test phase. The exception to this was Craig (column C, third graph). His performance deteriorated during the test until correct responding was at zero percent. Notations made on the data sheets during the test phase indicated that he consistently was touching the center (white space) of the stimulus sheets rather than consistently choosing the incorrect comparison stimulus.

Figure 9 presents these data in slightly different form. The three groups are presented on the horizontal axis. The data are plotted as the percentage of correct responses in the first half (first twenty trials) and the last half (second twenty trials) of the test phase. Thus, there are two data points for every child. The numbers 1 and 2 on the abscissa refer to the first half and the second half of the test phase respectively. The children are identified by letter. For example, in the normal group, "A" stands for Alex, "B"stands for Bobby, "C" stands for Claire, and "D" stands for Diane. Similarly, the letters ABCD identify the individual children in the other two groups.

In Figure 9, it is clear that all of the normal and retarded/language children improved in performance from the first to the last half of the testing period. For the normal children, the mean percentage of correct responding in the first half was 77.75%, and the mean for the second half was 95.5%. For the retarded/language children, the mean percentage of correct responding in the first half was 69.75%, and during the second half, 88%. For the retarded/no language group, mean correct responding during the first half was 46.25%, and during the second half, 39.25%.

Analyses of variance were done on the percentages of correct responding during first and second halves of the test phase. There was a significant difference among the groups on the percentage of correct responding in the first half of testing, $F(2, 9)=18.13$, $p<.0007$. There was also a significant difference among the groups on the percentage of correct responding in the last half of testing, $F(2, 9)= 18.68$, $p<.0006$. Newman-Keuls analyses showed no significant difference between the normal and the retarded/language groups in either the first or the last half of testing. Significant differences between the retarded/no language group and the normal ($p<.01$) and the retarded/language group ($p<.01$) were obtained in both the first and the last half of testing.

Supplementary Data

Mental age. To confirm that the mental age matching procedure had been effective in controlling for the influence of mental age per se on task performance, and that the language-deficit was the critical variable, the data were reanalyzed using mental age as the independent variable. This was done by dividing the children into two groups: High MA and Low MA. The children in the High MA group were Alex, Bobby, Allen, Beth, Andrew, and Barb. The children in the Low MA group were Claire, Diane, Carl, David, Craig, and Debbie. The division was made this way because the children's mental ages happened to fall into two relatively homogenous groups: those with mental ages 30-37 months and those with mental ages 14-20 months. The data for the number of trials to criterion in the conditional discrimination training are presented in Figure 3. No significant differences between the High and the Low MA groups were obtained on any dependent variable: trials to acquisition in the conditional discrimination training, $F(1, 10)=.85$, $p>.10$; the number of prompts used in training, $F(1,10)=.45$, $p>.10$; or the percentage of correct responses in the equivalence test, $F(1, 10)=.42$, $p>.10$.

Concrete stimuli. It seemed possible that the nonverbal children might have performed differently if larger or more concrete stimuli had been used. To assess this possibility, once Craig had completed the equivalence test, he was taught another series of conditional discriminations. The B, C, and E stimuli were novel visual stimuli (that had been used in another child's training as comparisons). The A stimulus (the sample) was a 6" yellow bottle of soap bubbles. The D stimulus (the other sample) was a pen flashlight. These materials were chosen as samples because they had been shown to be highly effective reinforcers. If the samples were highly salient, it seemed possible that the relationship between the comparisons paired with a particular sample would be more salient. The training procedure was identical to that used in the earlier training except that the samples and comparisons were physically quite dissimilar. The child was taught the tasks A-B, D-C, and A-E. Following this, the tasks were mixed. The bubbles and the penlight were used as reinforcers on the trials for which they also acted as sample stimuli. In the A-B, D-C, and A-E training, every correct response was consequated; and in the test phase the schedule was thinned to an approximate VR 3 for cooperation and good behavior.

Once he had met the mastery criterion on the mixed tasks, the test tasks were presented. On half of the trials, B was the sample, and C (incorrect) and E (correct) were the comparisons. On the other half of the trials, E was the sample, and B (incorrect) and C (correct) were the comparisons. A total of twenty test trials were presented. The data are presented in Figure 10. Tasks are identified along the horizontal axis; the percentages of correct unprompted responding within blocks of ten trials is presented along the vertical axis. "No responses" were not permitted; if one occurred, the trial was presented again. Craig's performance on the equivalence test was similar to the performances of the other children in the retarded/no language group on the original equivalence test. That is, his responding was at roughly chance level for this twenty trial equivalence test. His responding failed to show the establishment of a class of equivalent stimuli.

Patterns of fesponding during the test phase. Because of the way the test data had been collected, it was not possible to determine if the retarded/no language children were systematically making errors on some items (and responding correctly on others) or if responding on all trials was erratic. For example, it would have been possible for a child to obtain a overall score of 50% correct by consistently making the correct choice on B-E or E-B trials and choosing incorrectly on C-F and F-C trials. These additional data were collected to examine more closely responding during the test phase. The only modification of the original procedure made was that the child was required to select one of the two comparisons. If a "no response" occurred, or if the child touched the white space or perseverated in touching the sample, the experimenter said "No", frowned, briefly removed the materials, and re-presented the trial. Occasionally, if the child continued to fail to respond, the experimenter would prompt responding by raising the child's hand and holding it over the center of the stimulus sheet. This was not necessary more than once for each child during the test phase.

Debbie had earlier participated in training and testing (she was in the retarded/no language group). Randy (MA= 12 months, CA= 2 years, 5 months) had also received conditional discrimination training and equivalence testing although he was not an official project participant because of difficulties in obtaining appropriate mental age matches. His language/speech skills would have qualified him for membership in the retarded/no language group. Each of the children was retrained with the stimulus materials used in their earlier training.

The data are presented in Figure 11. The tasks are indicated along the horizontal axis; the data are graphed as the percentage of correct responses in blocks of ten trials. In the test phase, the data are presented both as the percentages of correct responding for each type of test stimulus and as the percentage correct for each consecutive block of ten trials. (Since the order of stimulus presentation was randomized during the test phase, the data on responding for each type of stimulus was not collected on consecutive trials.) The data indicate that for these two children, responding for each type of test stimulus was approximately 50%. In Debbie's case, responding for each of the test stimuli varied between 40% and 60% correct. This represents chance performance. Only 26 test trials were administered to Debbie because of time constraints. More test trials were administered to Randy; here is it clear that responding for each type of test stimulus was at 50% correct. In these two cases, then, chance responding in the test phase meant chance responding for each type of test stimulus.

Effects due to stimulus materials. It is possible, although unlikely, that the high levels of correct responding obtained in the test phase with the normal and retarded/language children might have been due to some characteristics of the stimulus sets rather than to the establishment of equivalence classes. That is, the sets may have been constructed in such a way that the children would have touched C in the presence of F, or B in the presence of E even without the previous training because of the physical similarity of the stimuli or prior extraexperimental history. Although this seemed an unlikely possibility because the experimental effect was obtained consistently across children in the two language-able groups (and because each stimulus set was individually constructed for each child), one set in which physical similarity may have played a role was identified. This was the set used to train and test Beth (Figure 12). The C and F stimuli in this case do share several common physical features.

To assess the role these physical similarities played in producing the high levels of correct responding obtained with Beth during the equivalence test, another child (Rachel, normally developing, experimentally naive, 18 months old) was given only the test trials and was told "Pick the right one". The results of this "test" may be seen in Figure 13. A total of twenty test trials were administered. Correct responding was at or about 50% for all types of test trials. This suggests that physical similarity is not an adequate explanation of the test results obtained with Beth.

CHAPTER IV

DISCUSSION

The results confirmed the prediction that the language-able children would perform better on the test of equivalence than the language-disabled children. In all instances, the language-able children performed better than chance during the equivalence test and in no instance did a language-disabled child perform better than chance during the equivalence test. The data supported the prediction that the retarded/no language children would perform poorly relative to the children in the other two groups on the equivalence test.

The retarded/no language children required significantly more trials to meet the mastery criterion in the conditional discrimination training portion of the project and required more prompts than the children in the other two groups. These data are consistent with reports in the literature on the acquisition of discriminations and conditional discriminations by retarded children with severely limited language abilities (e.g., Churchill, 1978; Lovaas, 1977; Routh, 1973). The performances of the normal children were consistent with the results obtained in similar studies (e.g., Gollin, 1964, 1965, 1966; Levin & Hammersmith, 1967). The data supported the prediction that the retarded/no language children would require more training and more assistance (prompts) during acquisition than the language-able children.

For the first time, in the research literature, a group of children who consistently fail to form equivalence classes has been identified.

(In previous studies, children occasionally failed to demonstrate the formation of equivalence classes; this is the first time that failure to form equivalence classes has been linked to a specific subject characteristic. In this project, the retarded/language children did succeed on the equivalence test, and the language-disabled children learned the conditional discriminations. Thus, the failure to demonstrate the formation of equivalence classes is a specific deficit not attributable solely to the presence of a handicapping condition (retardation) or to a general inability to learn. The position that stimulus equivalence is related to language is strengthened.

Failure to Demonstrate Stimulus Equivalence

Supplementary data supported the view that the poor performance on the equivalence task by the retarded/no language children was due to a specific deficit rather than to features of the training and testing or to mental age. For example, failure on the equivalence test was apparently not due to the use of abstract or unfamiliar stimuli. Even when three-dimensional, highly salient, materials were used in the conditional discrimination training (as with Craig), an equivalence class was not formed. The possibility that the excellent performance of the language-able children on the equivalence test was due to some characteristics of the stimulus materials, rather than to the preceding conditional discrimination training, was ruled out by using idiosyncratically constructed stimulus sets for each child. In one case in which it appeared that the characteristics of the stimuli used may have influenced equivalence test performance (Beth), supplementary data collected with another child (Rachel) supported the view that Beth's high rate of correct responding during the equivalence test was due to the conditional discrimination training and not to characteristics of the test stimuli.

The retarded/no language children could have performed at chance level on the equivalence test but still have consistently responded to one of the four types of test stimuli. This would have provided some evidence (albeit weak) of the formation of an equivalence class. Because of the way the data were collected, it was not possible to determine if this occurred. However, supplementary data collected with two children (Debbie and Randy) supported the view that, at least for some children, responding was poor across all four types of test trials. This is consistent with the view that no equivalence class was formed.

It appears likely that the failure of the nonverbal retarded children to demonstrate stimulus equivalence was due to failure to obtain symmetry (the functional substitutibility of the elements of the conditional discrimination) rather than to a failure to demonstrate generalized identity matching. This hypothesis is supported by informal assessments done with three children (Craig, Randy, and Andrew) at the conclusion of the study. The children were assessed for identity matching skills; one child was assessed for the presence of symmetrical relations. Each of the children was able to select an identical comparison from an array of five different items. The items used were colored blocks, small toys (two matching toy trucks), spoons, and snacks. After five training trials, on which correct matches were rewarded, three test trials were conducted. Two of the children, Craig and Randy, chose the identical comparison on all three trials. Andrew chose the correct comparison on two out of three trials. This indicates that the failure to demonstrate stimulus equivalence in these children was not due to a lack of generalized identity matching skills.

Craig was then assessed for the presence of symmetrical relations by taking one of his early original conditional discriminations (A-B) and making new stimulus sheets that used B as the sample stimulus and presented A and C as comparisons. These stimulus sheets were introduced after Craig had re-established a high rate of correct responding on the original discrimination and the reinforcement schedule had been thinned to an FR3. The new (B-A) stimulus sheets were presented on nonreinforced trials. Although A-B responding remained high, responding on the B-A trials was poor (2 out of 5 trials correct). This result differs from the results obtained in other studies (e.g., MacDonald, 1983) in which testing for symmetrical relations produced symmetry in children who had failed previously to demonstrate equivalence.

Thus, the failure to demonstrate stimulus equivalence, at least in some of these children, was apparently due to a failure to demonstrate symmetry. This suggests the possibility that training aimed at teaching symmetrical conditional relationships could, if successful, lead to improvements in the ability to form equivalence classes. In Craig's case, it might have been possible to obtain an improvement in responding on the B-A task had the testing been carried out longer. Direct teaching of symmetry may be necessary, however, for some children if extended testing does not produce it.

Acquisition of the Conditional Relationships

The analysis described thus far assumes that the children in the retarded/no language condition learned the conditional relationships. An alternative explanation of the failure to obtain equivalence is that the retarded/no language children learned a series of discrete stimulus-response pairs and learned nothing about the conditional relationships among the stimuli. For example, a child may have learned over the course of many trials to respond correctly on the A-B task and then to the A-E task, but not learned these as conditional tasks. Speaking loosely, the children may not have experienced the tasks as related at all. (Paul [1976] reports that this occurred in several of her young normal subjects in an auditory-visual matching task.)

Such a explanation seems quite possible in light of the problems in overly selective stimulus control repeatedly obtained with developmentally disabled children (Lovaas, Koegel, & Schriebman, 1979). It might account (in part) for the length of time needed for acquisition in the "mix" phases of the conditional discrimination training, although a systematic stimulus control analysis would be needed to determine the validity of the hypothesis. Alternatively, one could attempt to insure that the children learn conditionality by teaching large numbers of conditional discriminations. After the children had reached a high level of correct responding on a large number of conditional discriminations, a subset of the most recently trained stimuli could be used in an equivalence test. While this would not guarantee that the children had learned the conditional relationships, this training/testing would be a stronger test of the ability to form equivalence classes because it would enhance the likelihood that the conditional relations had been learned.

Improvement During Testing

All of the language-able children showed an improvement in test performance across blocks of test trials. This improvement in performance during testing has been seen in several studies (Fucini, 1982; Lazar, , Davis-Land, & Sanchez, 1984; Sidman, Kirk, & Willson-Morris, 1985). However, this study is one of the few in which the effect was seen in blocks of testing trials without conditional discrimination training interspersed between test trials (a similar result was obtained by Lazar et al., 1984). It has been suggested (Sidman et al., 1985) that this improvement occurs because the equivalence test itself provides a context in which the equivalence class is formed. The conditional discrimination training provides the necessary history and the introduction of testing trials provides the necessary context in which transitivity occurs and stimulus equivalence is demonstrated. The results of this project demonstrate that the interspersal of test and training trials is not necessary for the occurrence of improvement during testing and that improvement across the course of testing may be seen when testing trials are presented in a massed (as opposed to interspersed) format. That is, improvement across the course of testing does not require or is not dependent upon continued training on the conditional discrimination tasks.

The reasons for this improvement are not clear. One obvious possibility is that reinforcement for correct responding was occurring.

In this project, responses were neither explicitly rewarded or punished in the test phase, and any differential reinforcement occurring in the test phase was unprogrammed. (Cooperation and attention were consequated every three or four trials regardless of the correctness or incorrectness of the responses on the test.) One possible source of uncontrolled or unprogrammed reinforcement could have been subtle cues emitted by the experimenter (the Clever Hans effect, Sebeok, 1980). Although every effort was made to minimize any cues by the experimenter, the possibility of subtle differential reactions cannot be eliminated. However, this improvement during testing has been obtained repeatedly in studies in which automated equipment was used, and the possibility of differential feedback did not exist (e.g., Sidman et al., 1985). The consistency of the present data with the results of previous experiments makes the "Clever Hans" explanation of these results appear less likely.

The behavior of the normal and the retarded/language children during the test was consistent with the behavior of similarly aged children on other matching tasks (Levin & Maurer, 1969). That is, kindergarten and preschool children do not abandon a response when it fails to produce a reward but will perseverate in making their dominant response. In this case, the children did respond at times to the incorrect comparison stimuli in the initial portion of the test (as if to "check" on the contingencies that were in effect). In the absence of differential reinforcement, however, they persisted in making the responses that had been predisposed to favor as a result of the earlier conditional discrimination training. It is quite possible that the retarded/no language children made consistent choices in the test phase;

for example, a child may have consistently chosen the left side or the stimulus with corner angles (etc.). In any event, any consistent choices made by the retarded/no language children were not the choices they would have been predisposed to make if the conditional discrimination training had resulted in the formation of an equivalence class.

Generality of the Results

The generality of the failure to obtain stimulus equivalence with nonverbal children remains to be established. There is evidence that some retarded language-disabled children can do well on equivalence tests. Remington (personal communication, 1985) reports that nonverbal retarded children taught dictated name-picture (A-B) and picture-sign (B-C) associations produced the appropriate signs when given the dictated label (A-C), without having received any additional training. He suggested that receptive language may be an important variable in obtaining such transfer. In another project he reported (Remington, 1985), untrained picture-word associations emerged when the children comprehended the spoken object names but did not emerge when the children did not understand the spoken names.

The receptive language skills of the children in the present project were not formally assessed. All of the children responded to simple instructions, such as "Sit down" and to consequences such as "No!", but the extent to which they could comprehend labels and other conversation is not known. Since this study used only visual stimuli, the extent to which receptive language skills might have contributed to success or failure in establishing equivalence classes is not clear. Remington has suggested that the receptive label is important in mediating the association between the picture/object and its spoken name, a suggestion made by many other workers (e.g., Bialer, 1961).

In this case, it seems unlikely that all the children made up names for the stimuli in order to learn the conditional discriminations although it is possible that statements such as "green goes with red", referring to the colors of the stimuli, may have been used. There was no evidence for this in any of the the sessions. That is, none of the children made any comments that would indicate that they were using the color names to mediate responding. Further, since naming (receptive labeling) has been shown to be unnecessary for the establishment of equivalence classes, both when visual-only (Lazar et al., 1984) and when auditory-visual (Sidman et al., 1985) tasks are used. In other words, other studies have failed to demonstrate that naming is the critical variable in the formation of equivalence classes. Thus, a failure to name the stimulus should not have interfered with the children's ability to form equivalence classes. To put another way, the failure to use labels-as-mediators does not preclude the formation of equivalence classes (which, it is argued, is a more fundamental language skill). Because of this, it seems likely that the differences in performance between the children used in the two studies are due to other variables such as degree of mental or language impairment.

Equivalence and Language

These data do not allow us to determine the nature of the relationship between language skills and the ability to form equivalence classes. That is, the data do not tell us whether the ability to form equivalence classes is a product of language learning or a prerequisite to language learning.

It may be that children are able to form equivalence classes only after some degree of language acquisition has occurred. That is, the experiences that occur naturally in the course of learning to speak may also be experiences that result in the development of the generalized skill of equivalence class formation. This appears to be the view put forward by Hayes and Brownstein (1985). According to this position, stimulus equivalence is one example of a more general phenomenon-- the human propensity, given proper training, to respond to relationships between and among stimuli. Symmetry, for example, is in this view a generalized skill that develops after a consistent history of reinforcement has followed responding symmetrically to stimuli. That is, interaction with parents and others, in conjunction with the child's own labeling, may be critical experiences in the development of symmetry.

This analysis seems plausible if one considers that much of the behavior of parents with infants consists of teaching matching (e.g., Where is the doggie?, Point to the baby, Find the spoon). As the child begins to talk, the opportunity for the reinforcement of symmetry is created: the child says "doggie" in the appropriate context (and is reinforced) and points to the doggie when the parents say "doggie". In normal development, a great deal of such training occurs. This view that symmetry is dependent upon experience is also consistent with data from the psycholinguistic literature documenting asymmetry between children's production and comprehension of words (Nelson, Rescorla, Gruendel, & Benedict, 1978). Although initially a child's comprehension of a word and ability to utter it in the correct context are distinct, the comprehension/production disparity is often quickly resolved, and the child's use of the word becomes symmetrical (Gruendel, 1977). It is possible, then, that repeated hearing "doggie" in the presence of dogs and saying "doggie" in the same context and in the context of echoing parental productions leads to symmetry.

Conversely, the ability to form equivalence classes may be a language-learning prerequisite for normal human children. This appears to be consistent with the view put forward by Sidman (in press). According to this position, stimulus equivalence is an emergent property arising when humans learn related conditional discriminations. Sidman appears to argue that stimulus equivalence will appear in humans without any explicit training or particular history as a result of learning the appropriate conditional discriminations. Equivalence classes, then, in Sidman's view, are essential to the development of semantic and syntactical classes and the emergence of equivalence classes is not dependent on reinforcement history. Certainly one of the strongest pieces of evidence supporting this view is the ease with which most children acquire a first language across a wide range of linguistic environments. In addition, the success of the 25 month old in this project (Diane) indicates that very young human children are easily able to perform well on equivalence tasks (at least when they are presented visually).

The results of the present project are consistent with both positions. Apparently, in Sidman's view, the children may have failed to perform well on the equivalence test because they had failed to learn the conditional relationships among the stimuli. An alternative explanation might be that the children were so structurally damaged that that some important element or aspect of "humanness" is missing in them. It is difficult to account for these data otherwise as the children appeared to satisfy Sidman's two prerequisites for the formation of equivalence classes: being human and learning conditional discriminations.

The Sidman position does not suggest the possibility of improving children's skills in forming equivalence classes, although the Hayes/Brownstein position does. Since the emphasis in the Hayes/Brownstein position is on the organism's history with respect to the relationships between stimuli and not on the emergent character of stimulus equivalence (Sidman's view), the former view would predict that provision of the proper history should result in the development of symmetry and, eventually, the formation of equivalence classes.

The relationship between the ability to form equivalence classes and language might be clarified through longitudinal studies of normal babies. Investigations of equivalence test performance before and following the appearance of the child's first words and at other regular intervals could identify the point in language acquisition that the child is able to form equivalence classes. If, for example, performance on an equivalence test is poor before the child has acquired a vocabulary of several words, the hypotheses that the ability to form equivalence classes requires a reinforcement history to appear or is a product of language acquisition would be strengthened. Good performance on an equivalence test prior to the acquisition of the child's first words would strengthen the argument that the ability to form equivalence classes is a precursor of language acquisition.

Another potentially fruitful approach would be to examine the equivalence test performance of aphasic children with normal performance (nonverbal) IQs. Such research may help to resolve the continuing controversy over of whether such children suffer a fundamental inability to use symbols (Christpoulou & Bonvillian, 1985).

Limitations of the Present Study

The limitations of the correlational approach have already been discussed. That is, it is not possible to determine the nature of the relationship between stimulus equivalence and language from these data. In addition, it appears that the experiment was not a completely pure test of the hypothesis as it is obvious that the retarded/no language children suffered greater impairments than simply a lack of language. Each of the retarded/no language children showed deficits in all areas of social and adaptive functioning. All of the children in the retarded/no language group were students in classes for the severely and profoundly retarded while the children in the retaroed/language group were students in classes for the mild and moderately retarded. The differences among the children in social, play, self-help, and intellectual behaviors indicated that the mental-age matching procedure did not succeed in equating the children in all areas except for language skills. The mental age matching procedure is known to be most effective when the mental retardation is nonpathological in origin (Achenbach, 1982). That is, MA matching is most successful in equating the skills or normal children and children whose retardation is due to socio-cultural (nonorganic) causes.

These differences between the children were an artifact of the selection process. The children were chosen for participation based on their expressive language skills and parental willingness to allow participation. Although some students in the severely and profoundly retarded class used some functional words or signs, parental permission for participation was not given for these children and so they were not included. Thus, the possibility that equivalence test performance differences between the retarded/language and normal children and the retarded/no language children were due to impairments other than language dysfunction cannot be ruled out.

Summary

These data suggest that it may be worthwhile to pursue additional research on the relationship between language and stimulus equivalence in the hopes of providing tools for the remediation of language and generalization deficits in developmentally disabled populations. Some of the supplementary data collected in the project have shown that a failure to show symmetrical relations may have been responsible for the failure to establish equivalence classes in at least some cases.

Another possible contributor to the failure to demonstrate the formation of equivalence classes could be the children's failure to learn the conditional relationships. These suggest that two logical steps for future research would be to attempt to teach symmetry to children who fail to show them after testing and to insure that the children learn the conditional relations by extensive training before testing for equivalence.

The psycholinguistic literature has shown an increasing appreciation for the contributions of environmental context and history to language development (cf. Bruner, 1981; Furrow & Nelson, 1984; Holzman, 1984). Continued research on the relationship between stimulus equivalence and language will enhance our ability to analyze semantic (Segal, 1975; Sidman, in press) and syntactical (Lazar, 1977) development and may provide a bridge between the experimental analysis of behavior and psycholinguistics.

BIBLIOGRAPHY

Achenbach, T. M. (1969). Conservation of illusion distorted identity: Its relation to CA and MA in normals and retardates. Child Development, 40, 663-679.

Achenbach, T. M. (1970). A comparison of Stanford-Binet performances of retarded and nonretarded persons matched for MA and sex. American Journal of Mental Deficiency, 74, 488-494.

Achenbach, T. M. (1982). Developmental psychopathology. New York: John Wiley.

Anderson, S. R. & Spradlin, J. E. (1980). The generalized effects of productive labeling involving common object classes. JASH, 5, 143-157.

Bialer, I. (1961). Primary and secondary stimulus generalization as related to intelligence level. Journal of Experimental Psychology, 62, 395-402.

Bronowski, J. & Bellugi, U. (1970). Language, name and concept. Science, 168, 669-673.

Bruner, J. (1981). The social context of language acquisition. Language and Communication, 1, 155-178.

Carter, D. E. & Werner, T. J. (1978). Complex learning and information processing by pigeons: A critical analysis. Journal of the Experimental Analysis of Behavior, 29 565-601.

Christopoulou, C. & Bonvillian, J. M. (1985). Sign language, pantomime, and gestural processing in aphasic persons. Journal of Communication Disorders, 18, 1-20.

Chomsky, N. (1969). Human language and other semiotic systems. Semiotica, 25, 31-44.

Churchill, D. (1978). Language of autistic children. New York: John Wiley.

Clark, H. H. & Clark, E. V. (1978). Psychology and language: An introduction to psycholinguistics. New York: Harcourt, Brace, Jovanovich, Inc.

Curtiss, S. (1981). Dissociations between language and cognition: Cases and implications. Journal of Autism and Developmental Disabilities, 11, 15-30.

D'Amato, M. R., Salmon, D. P. & Columbo, M. (1985) Extent and limits of the matching concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 35-51.

Dixon, M. (1976). Teaching conceptual classes with receptive label training, *Acta Symbolica*, *9*, 17-35.

Dixon, M. & Spradlin, J. E. (1976). Establishing stimulus equivalences among retarded adolescents. *Journal of Experimental Child Psychology*, *21*, 144-164.

Edwards, C.A., Jazielo, J. A., & Zentall, T. R. (1983). "Same/different" symbol use by pigeons. *Animal Learning and Behavior*, *11*, 349-355.

Ellis, A. (in press). *Progress in the psychology of language. Volume 2*. Hillsdale, New Jersey: Lawrence Erlbaum.

Ferguson, G. (1976). *Statistical analysis in psychology and education*. New York: McGraw-Hill.

Fouts, R. F., Chown, W., & Goodin, L. T. (1976). Transfer of signed responses in American Sign Language from vocal English stimuli to physical object stimuli by a chimpanzee (*Pan*). *Learning and Motivation*, *7*, 458-475.

Fucini, A. (1982). Stimulus control of class membership. Unpublished doctoral dissertation, Northeastern University.

Furrow, D. & Nelson, K. (1984). Environmental correlates of individual differences in language acquisition. *Journal of Child Language*, *11*, 523-534.

Gardner, R. A. & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, *165*, 664-672.

Gast, D. vanBiervlet, A. & Spradlin, J. (1979). Teaching number-word equivalences: A study of transfer. *American Journal of Mental Deficiency*, *83*, 524-527.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, *49*, 988-998.

Gollin, E. G. (1964). Reversal learning and conditional discriminations in children. *Journal of Comparative and Physiological Psychology*, *53*, 441-445.

Gollin, E. G. (1965). Factors affecting conditional discrimination in children. *Journal of Comparative and Physiological Psychology*, *60*, 422-427.

Gollin, E. G. (1966). Solution of conditional discrimination problems by young children. Journal of Comparative and Physiological Psychology, 62, 454-456.

Gruendel, J. (1977). Referential extension in early language development. Child Development, 48, 1567-1576.

Harris, S. (1975). Teaching language to nonverbal children- with special emphasis on problems of generalization. Psychological Bulletin, 82, 565-580.

Hayes, S. C. & Brownstein, A. (1985). Verbal behavior, equivalence classes and rules: New definitions, data and directions. Paper presented at the Association for Behavior Analysis Annual Meeting, May 1985.

Hayes, S. C., Brownstein, A., Devany, J. M., Kohlenberg, B. & Shelby, J. (1985). Stimulus equivalence and the acquisition of meaning. Paper presented at the Association for Behavior Analysis Annual Meeting, May 1985.

Herman, L. M. & Thompson, R. K. (1982). Symbolic, identity and probe-delayed matching of sounds by the bottle-nosed dolphin. Animal Learning and Behavior, 10, 22-34.

Holzmann, M. (1984). Evidence for a reciprocal model of language development. Journal of Psycholinguistic Research, 13, 119-146.

Jenkins, J. J. (1965). Mediation theory and grammatical behavior. In S. Rosenberg (Ed.), Directions in psycholinguistics. (pp. 66-96). New York: MacMillan.

Karlan, G. R. (1976). An investigation of auditory stimulus equivalence in the context of natural word classes with preschool children. Unpublished doctoral dissertation, University of Kansas.

Kendall, S. B. (1983). Tests for mediated transfer in pigeons. Psychological Record, 33, 245-256.

Lashley, K. (1938). Conditional reactions in the rat. Journal of Psychology, 6, 311-324.

Lazar, R. (1977). Extending sequence-class membership with matching-to-sample. Journal of the Experimental Analysis of Behavior, 27, 381-392.

Lazar, R., Davis-Lang, D. & Sanchez, L. (1984). The formation of visual stimulus equivalences in children. Journal of the Experimental Analysis of Behavior, 41, 251--266.

Lenneberg, E. (1969). A word between us. In J. D. Roslansky (Ed.) Communication (pp. 111-131). Amsterdam: North Holland.

Levin, G. R. & Hammersmith, D. R. (1967). Procedures and instructions in children's matching-to-sample. Psychonomic Science, 8, 429-430.

Levin, G. R. & Maurer, D. M. (1969). The solution process in children's matching to sample. Developmental Psychology, 1, 679-690.

Lieber, J. (1982). Language and its analogues: Functional and homological characterizations in cognitive psychology. Language and Communication, 2, 219-225.

Lovaas, O. I. (1977). The autistic child. New York: Irvington.

Lovaas, O. I., Koegel, R., & Schriebman, L. (1979). Stimulus overselectivity in autism: A review of research. Psychological Bulletin, 86, 1236-1254.

MacDonald, R. F. (1983). Match-to-sample performance in the context of group teaching. Unpublished doctoral dissertation, University of Kansas.

Mackay, H. A. & Sidman, M. (1985). Teaching new behaviors via equivalence relations. In P. Brooks, R. Sperber, & C. McCauley (Eds.) Learning and cognition in the mentally retarded. (pp. 493-513). Hillsdale, New Jersey: Lawrence Erlbaum.

Muncer, S.J. & Ettinger, G. (1984). In and behind a Lenneberg paradigm. Journal of Psycholinguistic Research, 13, 57-68.

Nelson, K. (1974). Concept, word and sentence: Interrelatons in acquisition and development. Psychological Review, 81, 267-285.

Nelson, K., Rescorla, L., Gruendel, J., & Benedict, H. (1978). Early lexicons: What do they mean? Child Development, 49, 960-968.

Nissen, H. (1951). Analysis of complex conditional reaction in the chimpanzee. Journal of Comparative and Physiological Psychology, 7, 449-516.

Pate, J. L. & Rumbaugh, D.M. (1983). The language-like behavior of Lana chimpanzee: Is it merely discrimination and paired associate learning? Animal Learning and Behavior, 11, 134-138.

Patterson, F. (1978). The gestures of a gorilla: Language acquisition in another pongid. Brain & Language, 5, 56-71.

Paul, L. (1976). Stimulus class learning by elimination. Unpublished master's thesis. University of Kansas.

Pepperberg, I. (1983). Cognition in the African Grey parrot: Preliminary evidence for auditory-vocal comprehension of the class concept. Animal Learning and Behavior, 11, 179-185.

Premack, D. (1983). The codes of man and beasts. Behavioral and Brain Sciences, 6, 125-167.

Remington, R. (1985). Is signing a substitute for speech with retarded children? Paper presented at the University of North Carolina (Greensboro), April 25, 1985.

Ristau, C. A. & Robbins, D. (1982). Cognitive aspects of ape language experiments. In D. R. Griffin (Ed.) Animal Mind-Human Mind. (pp. 299-331). New York: Springer-Verlag.

Routh, D. (1973). The experimental psychology of mental retardation. Chicago: Aldine .

Sebeok, T. (1980). Looking in the destination for what should have been sought in the source. In T. Sebeok and J. Umiker-Sebeok (Eds.) Speaking of apes. (pp. 407-428). New York: Plenum.

Segal, E. F. (1975). Psycholinguistics discovers the operant: A review of Roger Brown's A First Language. Journal of the Experimental Analysis of Behavior, 23, 149-158.

Shvartz, L. A. (1960). Uslovnyye refleksy na slovenyye razdrazhiteli. Voprosy psikhologii, 6, 86-98. Described in D. I. Slobin (1971), Psycholinguistics. (pp. 84-85). Glenview, Illinois: Scott Foresman.

Sidman, M. (1971). Reading and auditory-visual equivalences. Journal of Speech and Hearing Research, 14, 5-13.

Sidman, M. (in press). Functional analyses of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.) Units of analysis and integration of behavior. Hillsdale, New Jersey: Lawrence Erlbaum.

Sidman, M. & Cresson, O. (1973). Reading and cross-modal transfer of stimulus equivalences in severe retardation. American Journal of Mental Deficiency, 77, 515-523.

Sidman, M., Cresson, O., & Willson-Morris, M. (1974). Acquisition of matching-to-sample via mediated transfer. Journal of the Experimental Analysis of Behavior, 22, 261-273.

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons and children. Journal of the Experimental Analysis of Behavior, 37, 23-44.

Sidman, M. & Tailby, W. (1982). Conditional discrimination versus matching to sample: An expansion of the testing paradigm. Journal of the Experimental Analysis of Behavior, 37, 5-22.

Sidman, M., Kirk, B.,& Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. Journal of the Experimental Analysis of Behavior, 43, 21-42.

Skinner, B.F. (1957). Verbal Behavior. NY: Appleton-Century-Crofts.

Skinner, B. F. (1974). About Behaviorism. NY: Vintage Books.

Slobin, D. I. (1971). Psycholinguistics. Glenview, Illinois: Scott Foresman.

Spradlin, J. E., Cotter, V. W. & Baxley, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. American Journal of Mental Deficiency, 77, 556-566.

Spradlin, J. E. & Dixon, M. (1976). Establishing a conditional discrimination without direct training: Stimulus classes and labels. American Journal of Mental Deficiency, 80, 555-561.

Stokes, T. & Baer, D. M. (1977). An implicit technology of generalization. Journal of Applied Behavior Analysis, 10, 349-368.

Turney, T. H. (1982). The associations of visual concepts and imitative vocalizations in the mynah. Bulletin of the Psychonomic Society, 19, 59-62.

vanBiervliet, A. (1977). Establishing words and objects as functionally equivalent through manual sign training. American Journal of Mental Deficiency, 82, 178-186.

Wetherby, B., Karlan, G. R., & Spradlin, J. E. (1984). The development of derived stimulus relations through training in arbitrary matching sequences. Journal of the Experimental Analysis of Behavior, 40, 69-78.

APPENDIX A

FIGURES

Figure 1. The stimulus equivalence paradigm (cf. Sidman, Cresson, & Willson-Morris, 1974).

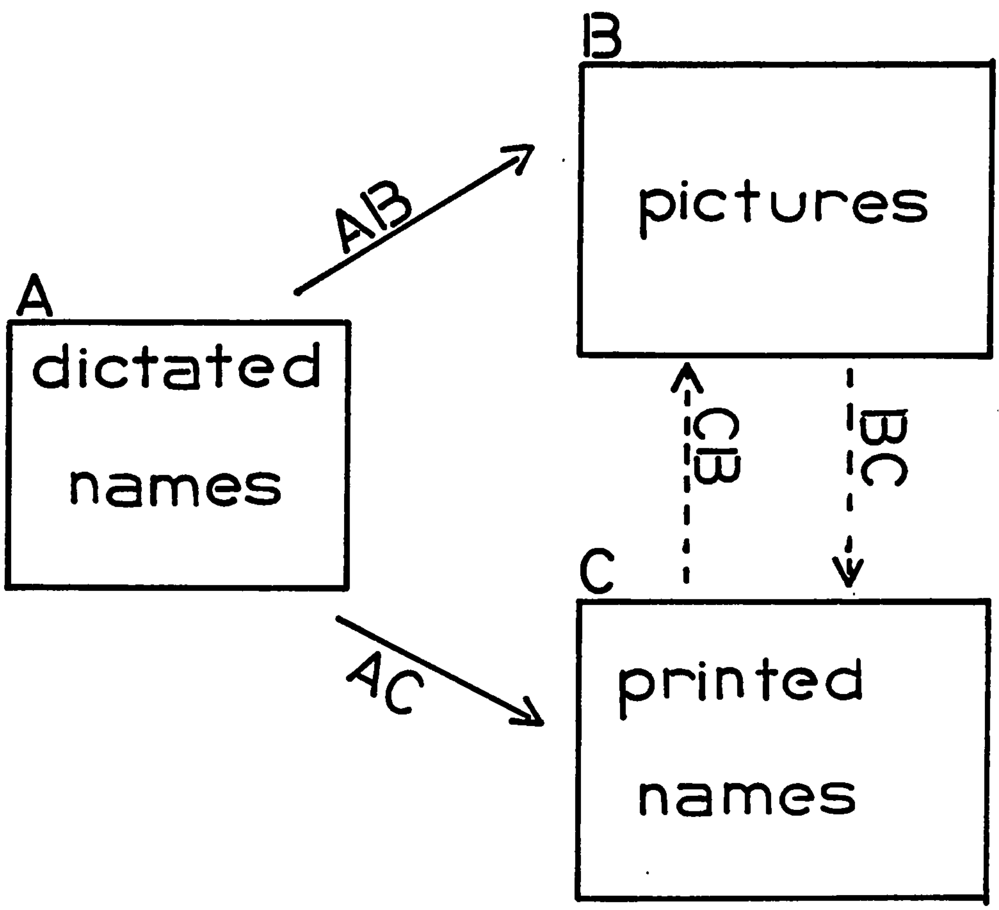


Figure 2. The stimulus figures used in this experiment.

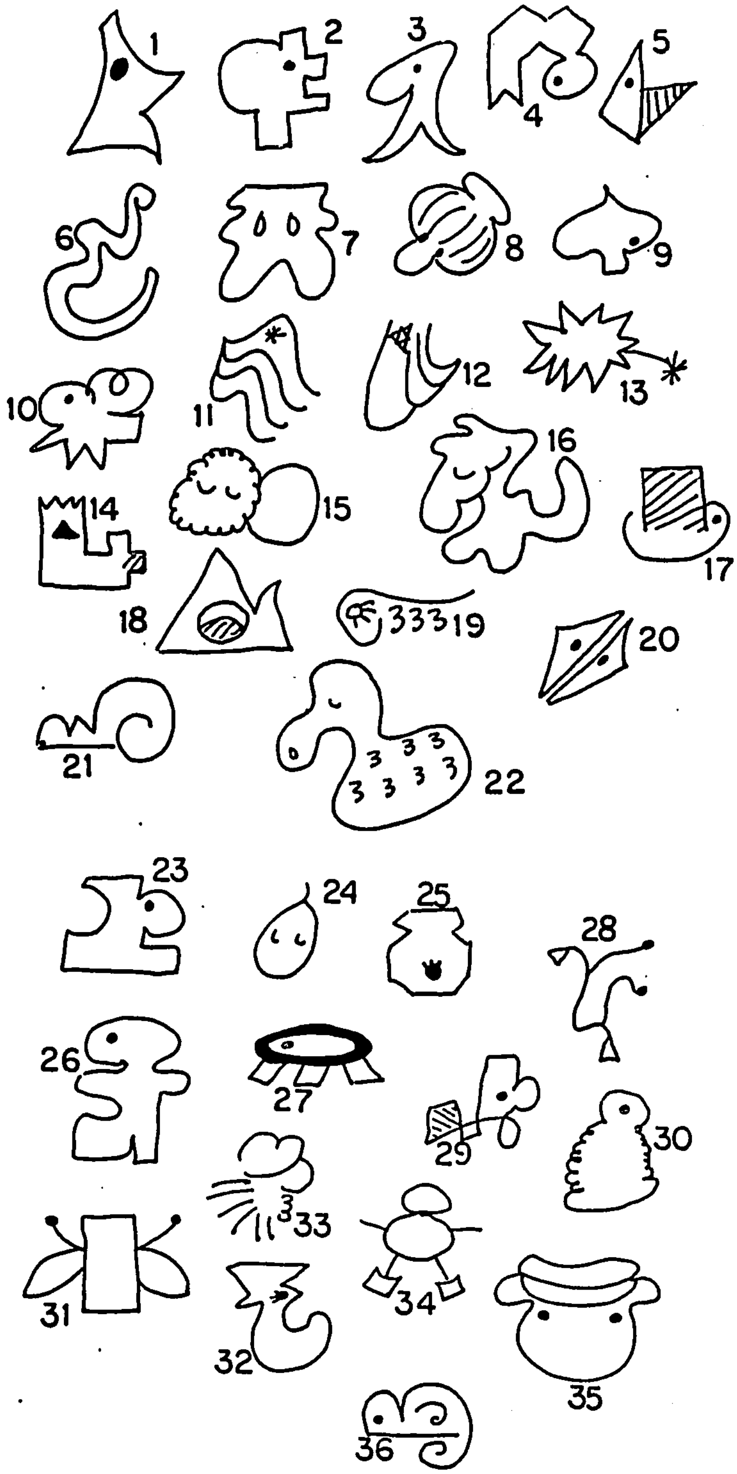


Figure 3. The number of trials to mastery criterion during the conditional discrimination training. The data are presented by group and by mental age. The number of trials is indicated along the vertical axis; groups and mental ages are indicated aong the horizontal axis.

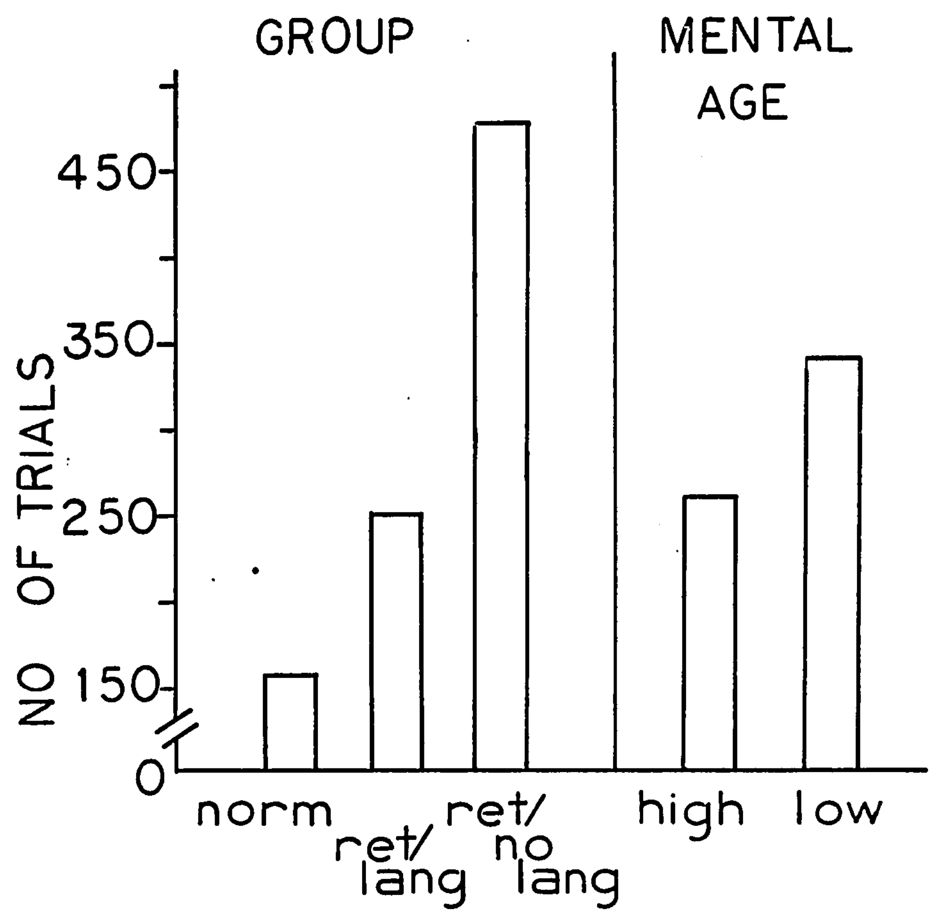


Figure 4. Individual training and testing data for Alex, Allen and Andrew. The data are graphed as the percentage of correct unprompted trials (vertical axis) across blocks of ten trials (horizontal axis). The numbers within each graph (1, 2, 3, and so on) represent the training phases. Number 1 refers to the A-B training, number 2 to the D-C training, 3 to the A-B D-C mix, 4 to the A-E training, 5 to the D-F training, 6 to the A-E D-F mix, and 7 to the final mix of all four conditional tasks. The 40 trial test phase is indicated on each graph.

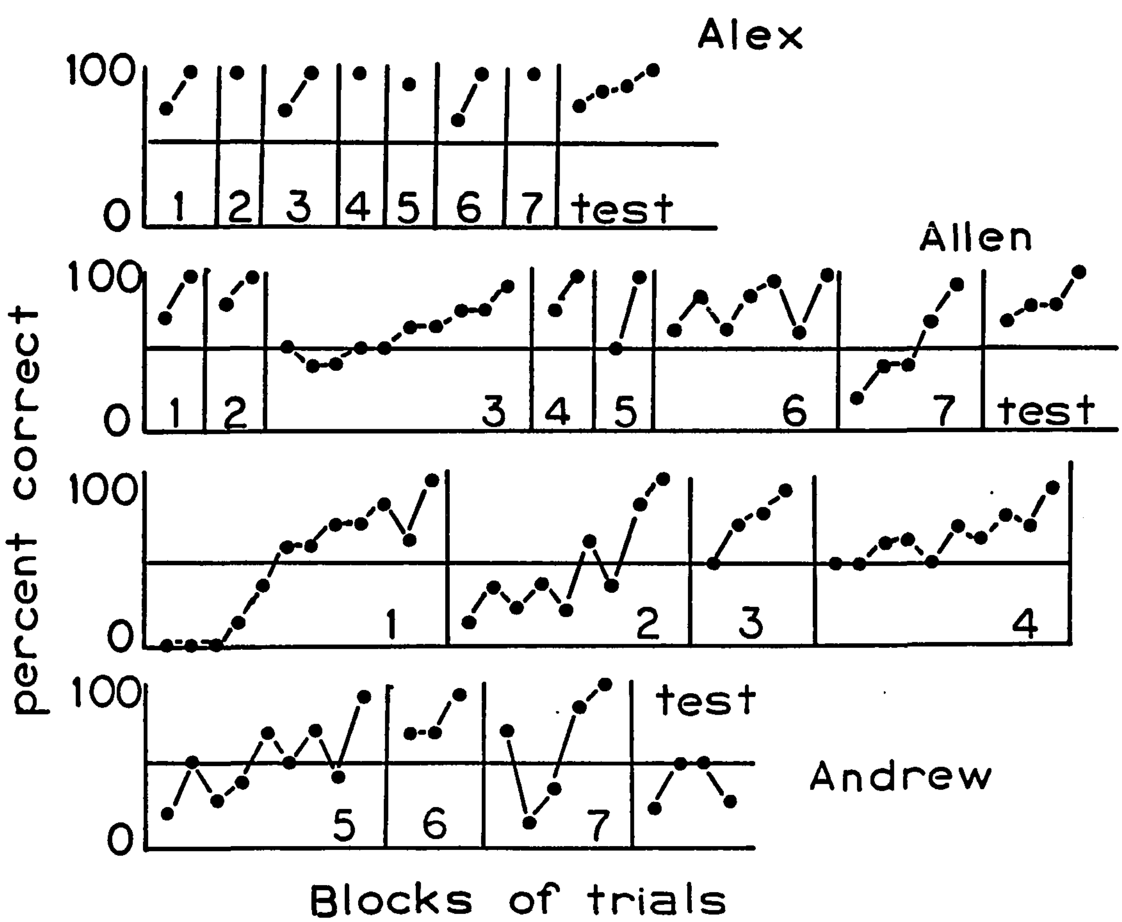


Figure 5. Individual training and testing data for Bobby, Beth and Barb.

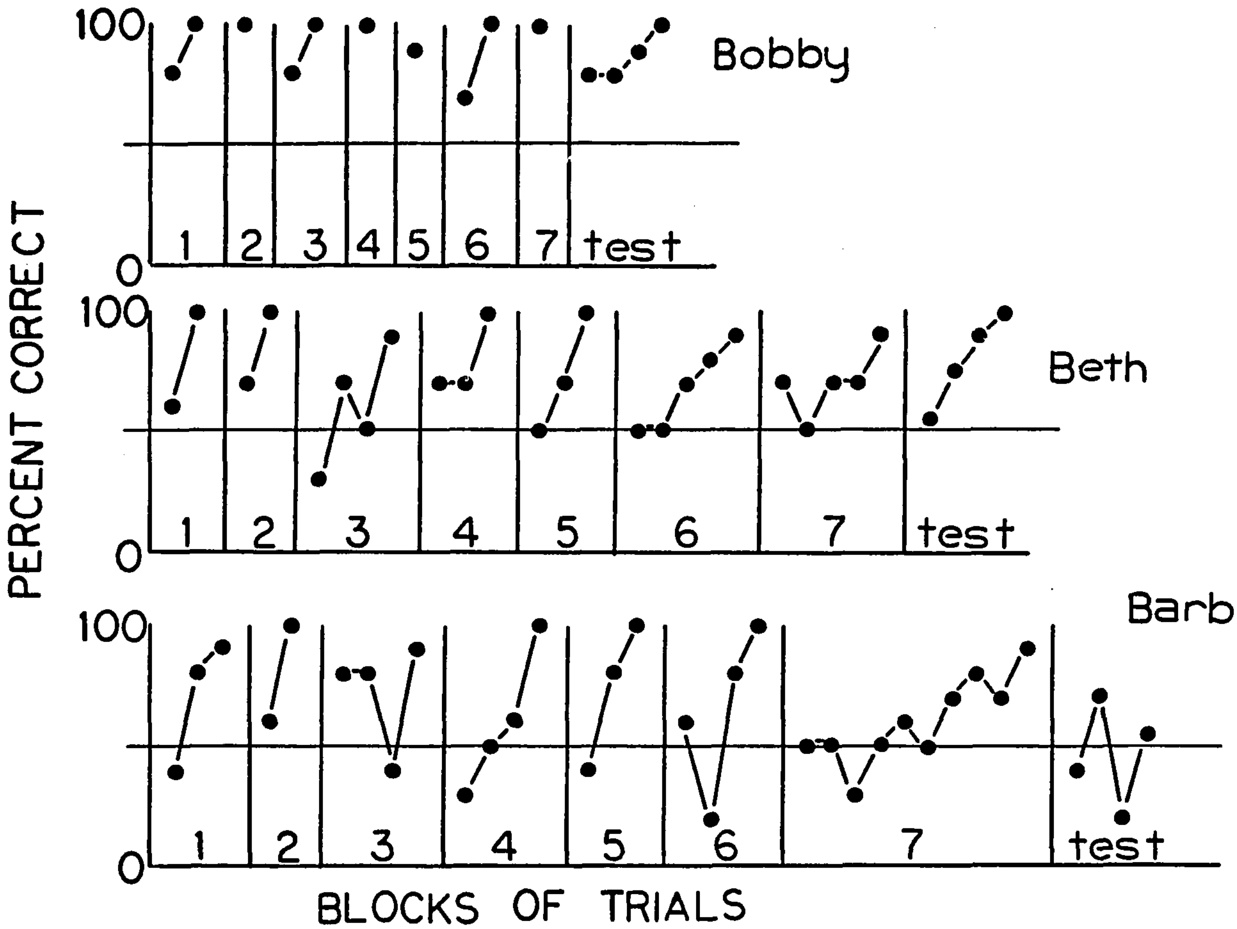


Figure 6. Individual training and testing data for Claire, Carl and Craig.

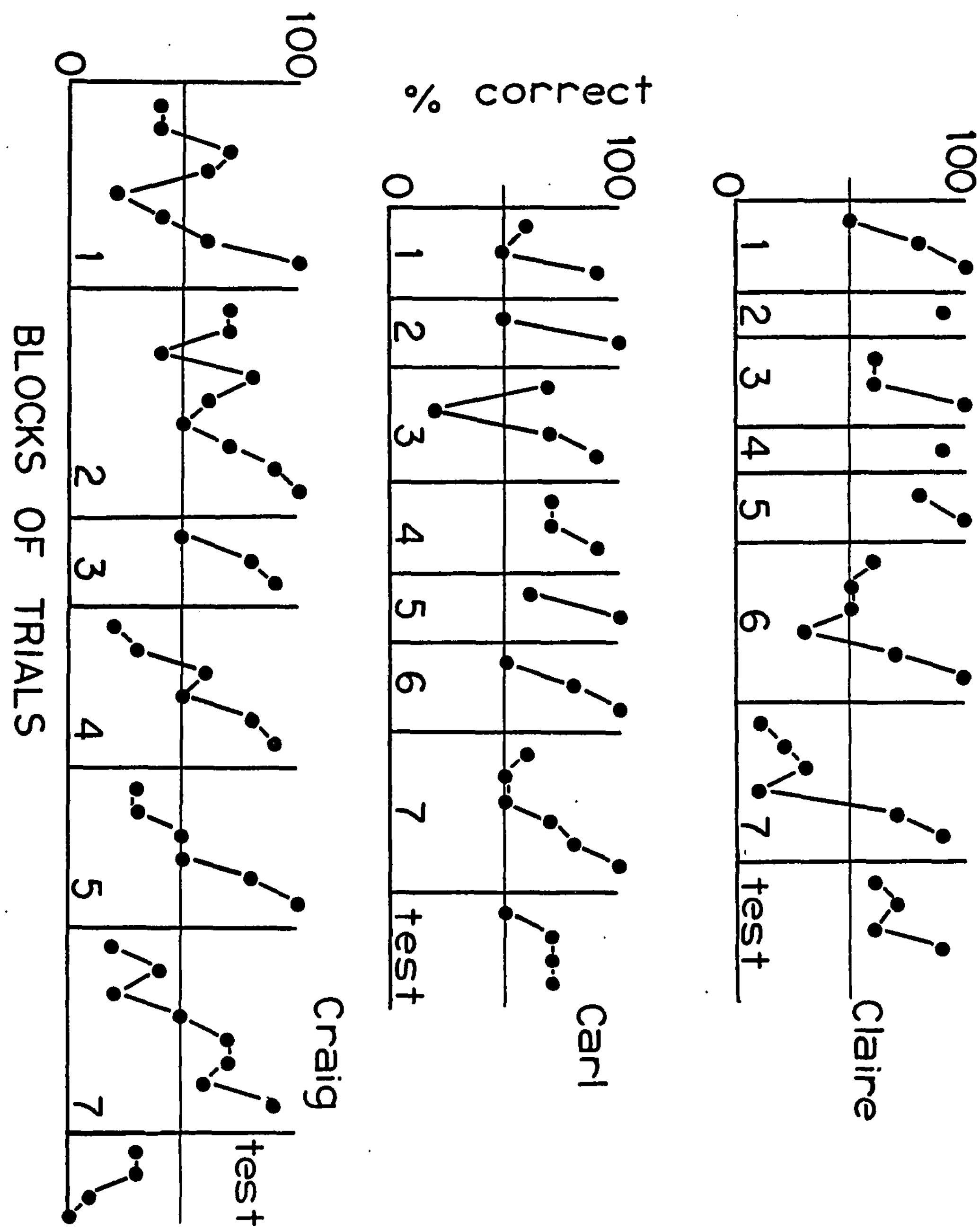


Figure 7. Individual training and testing data for Diane, David and Debbie.

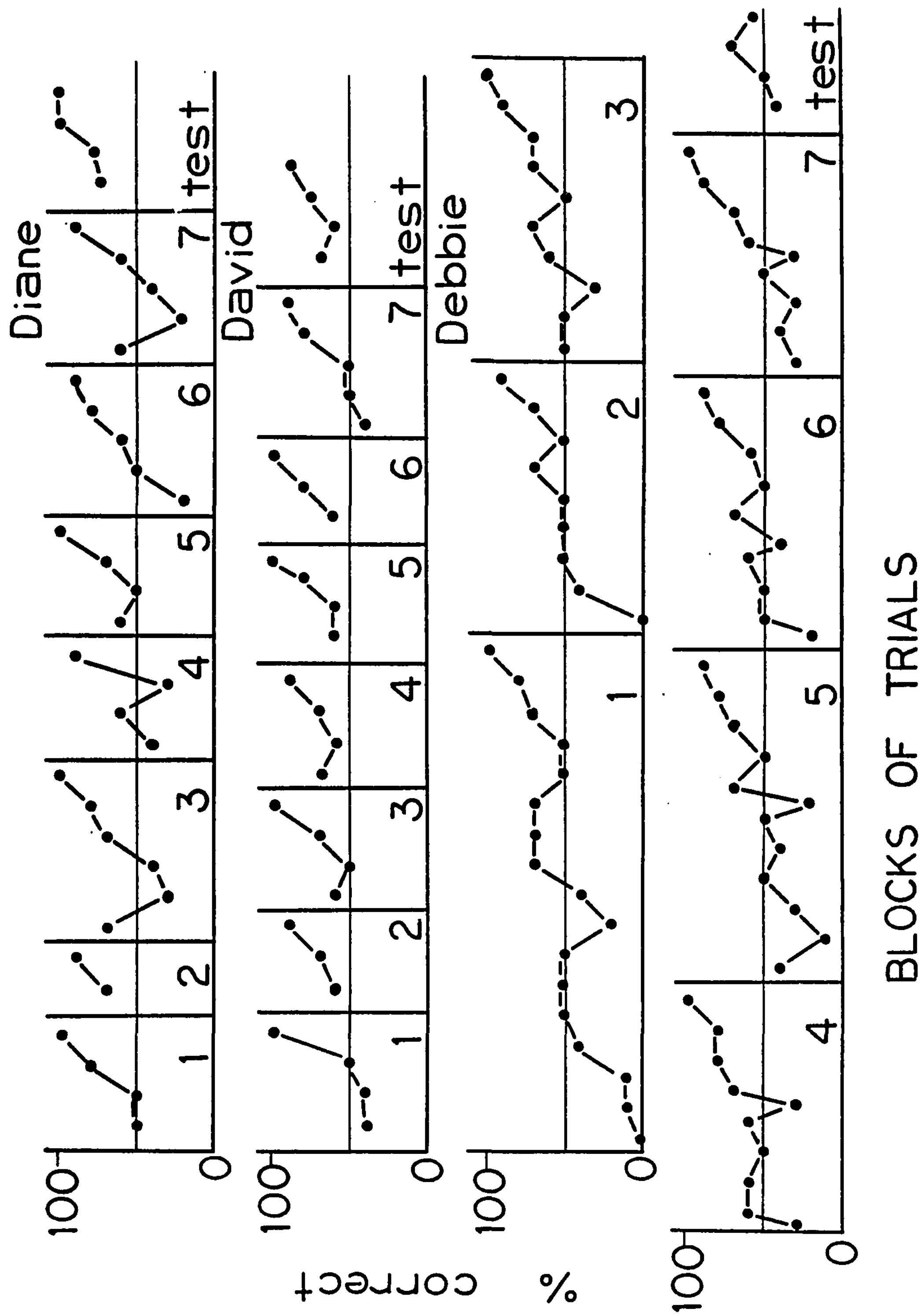


Figure 8. Responding during the test phase. Each graph presents the data for one child. Each column of graphs presents the data for those children matched for mental age. Within each column, the top graph presents the data for the normal child, the middle graph presents the data from the retarded/language child and the bottom graph presents the data from the retarded/no language child. The data (dots) are presented as the percentage correct (of all responses attempted) across blocks of ten trials. The columns within each graph present the number of no responses that occurred within that block of ten trials.

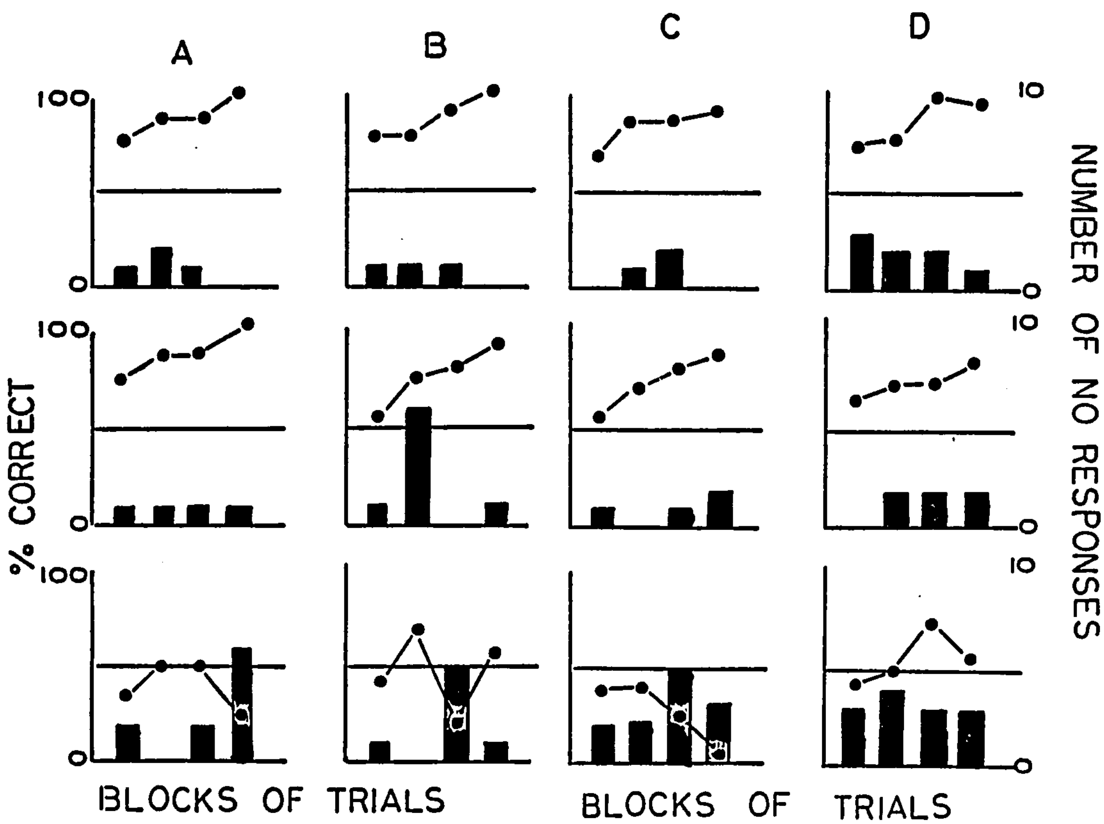


Figure 9. Testing data. The data are graphed as the percentage of correct responses (out of responses made) during the first twenty trials and last twenty trials of the test phase. The percentage of correct responses is plotted on the vertical axis; the test halves and groups are identified on the horizontal axis. Individual children within groups are identified by the letter corresponding to their pseudonym.

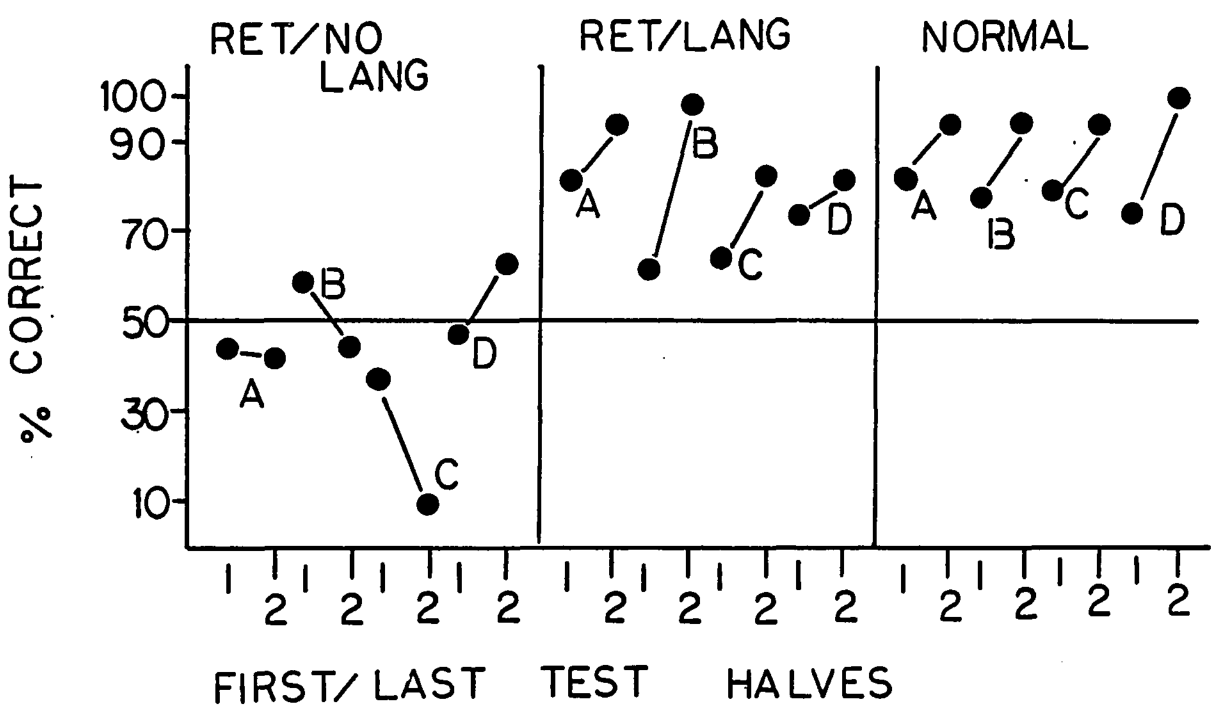


Figure 10. Supplementary data collected with Craig. The data are presented as the percentage of correct responses across blocks of ten trials. The conditional discriminations and test phase are indicated in the figure.

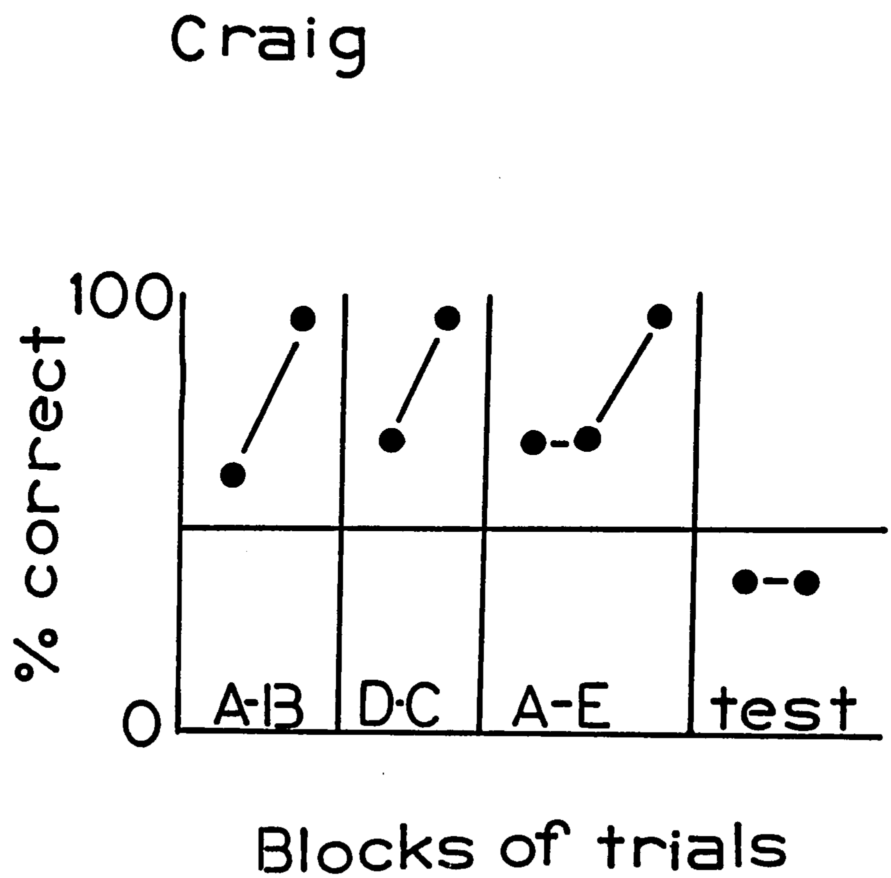


Figure 11. Retraining and testing data collected with Debbie and Randy. The conditional discriminations (retrained) are indicated by the numbers in the figures. Responding during the test phase is graphed both as the number of correct responses made in the presence of specific test stimuli (test items) and as the number of correct responses made during consecutive blocks of ten trials (blocks). The cross next to the final data point in the "block" test session for Debbie indicates that this was not a complete block of ten trials.

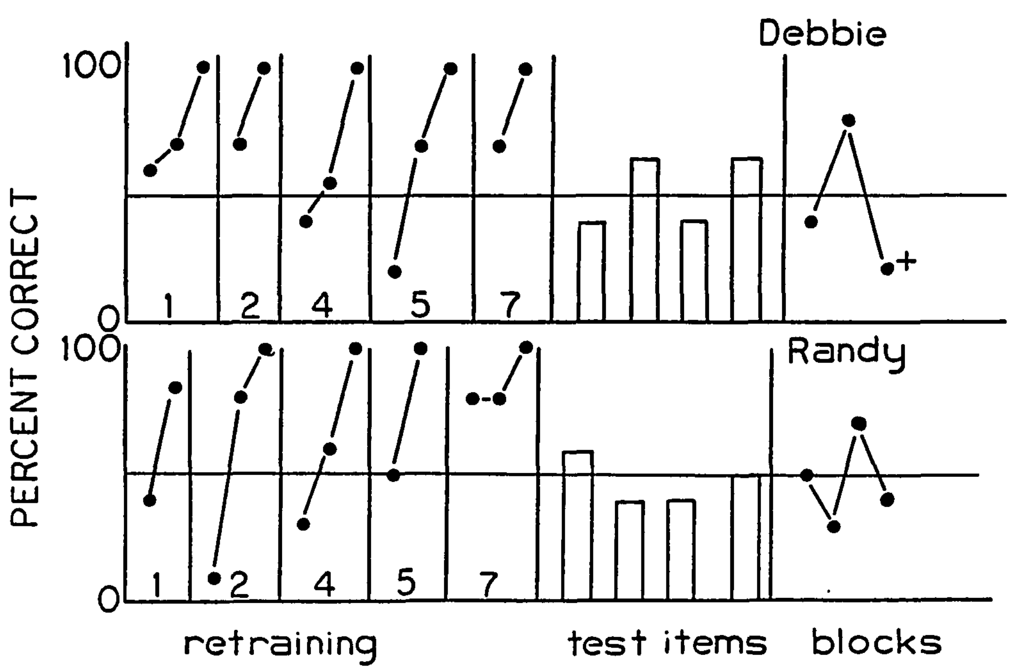


Figure 12. Stimulus set used for training and testing Beth. The trained relations are indicated by the heavy black arrows; the tested relations are indicated by the slender arrows.

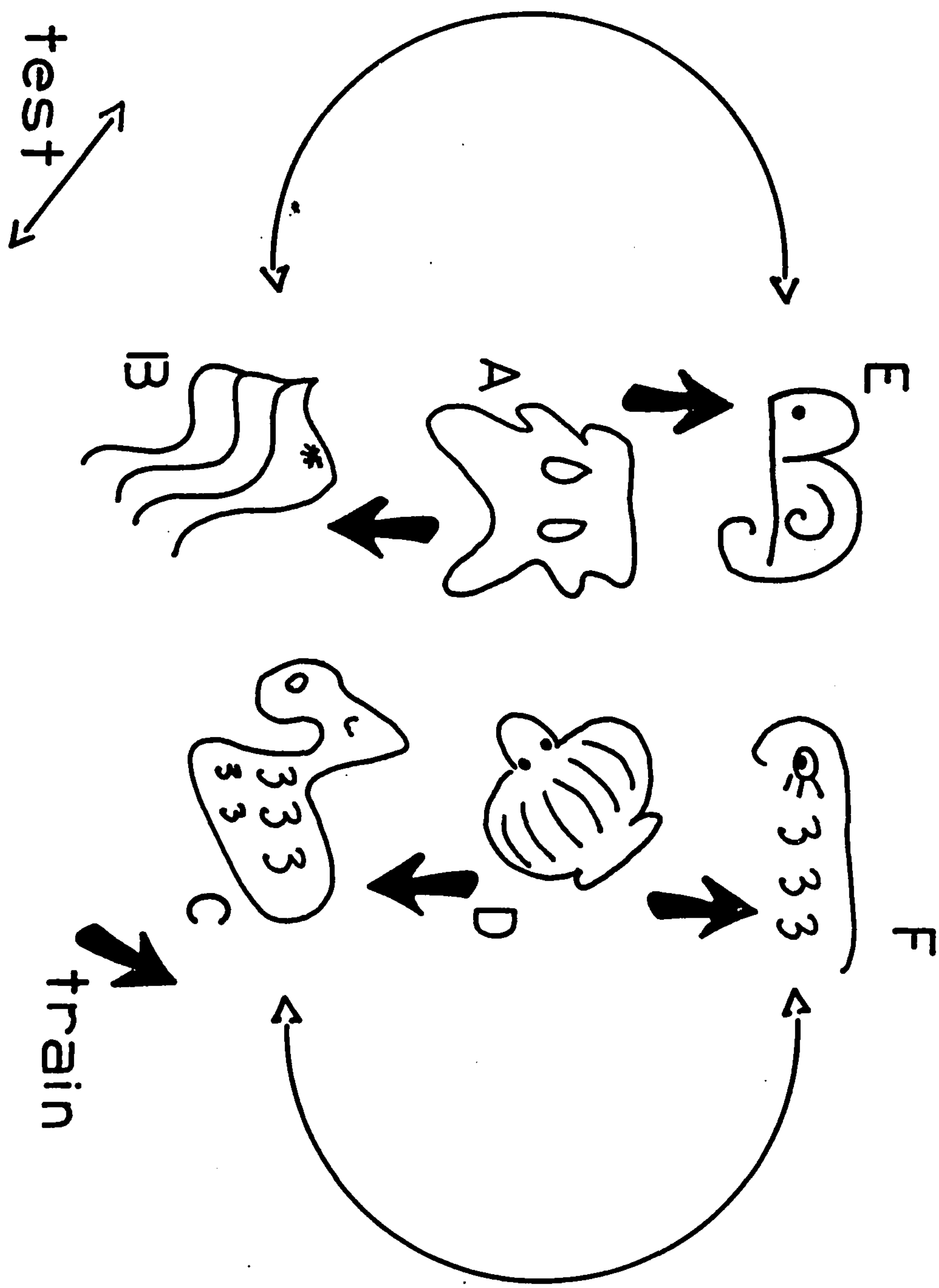
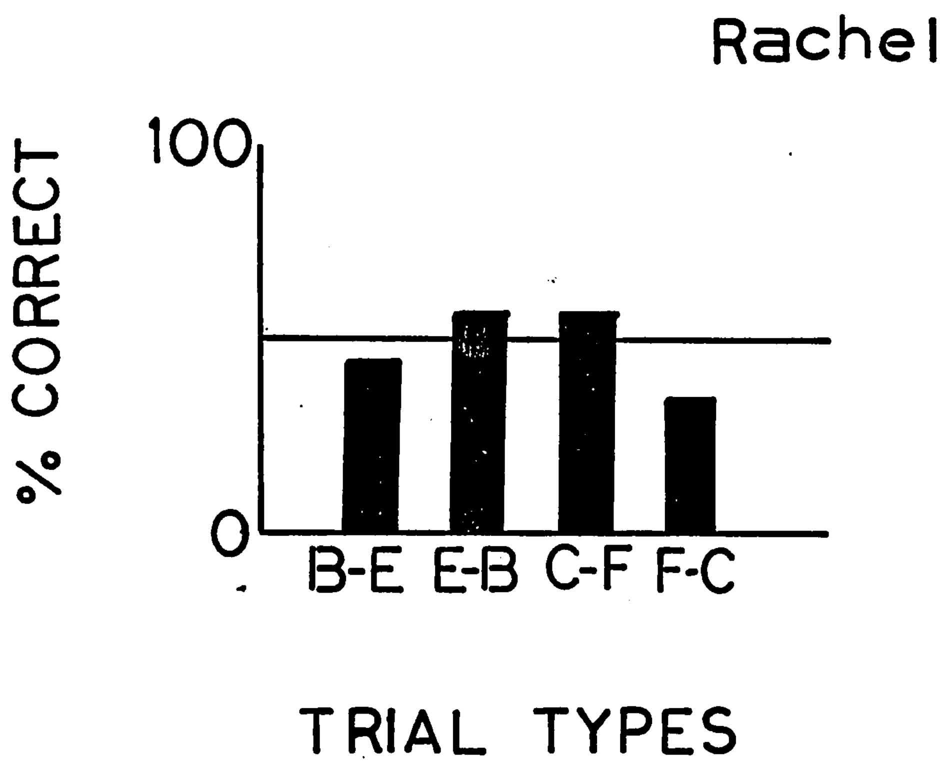


Figure 13. Rachel's "test". The data are graphed as the precentage of correct responses made in the presence of each of four trial types. The stimuli used were the test stimuli used with Beth.



APPENDIX B

TABLES

Table 1

Subject Characteristics

	Mental Age	Chronological age
	Normal Children	
Alex	30 months	2-6
Bobby	37 months	2-11
Claire	20 months	2-6
Diane	19 months	2-1
	Retarded/Language Children	
Allen	31 months	3-7
Beth	36 months	4-4
Carl	20 months	3-3
David	19 months	2-8
	Retarded/No Language Children	
Andrew	30 months	4-1
Barb	36 months	4-4
Craig	18 months	4-4
Debbie	14 months	2-7

Table 2

Reliability Data

Date	Number of Trials Observed	Percent Agreement
February 5	67	88
February 12	60	100
Feoruary 27	150	96
March 6	60	100
March 14	200	100
March 18	120	100