Analysis and Intervention in Developmental Disabilities, Vol. 1, pp. 259–297, 1981
Printed in the USA.
0270–4684/81/030259–39$02.00/0

Teaching Complex Language to Autistic Children

Patricia J. Krantz, Stanley Zalenski, Laura J. Hall, Edward C. Fenske and Lynn E. McClannahan

Princeton Child Development Institute

Although there is a growing technology for teaching autistic children to engage in verbal imitation, to begin with nouns, verbs and pronouns, and to construct simple sentences, there is a continuing need for investigations of procedures that may be used to teach additional language skills. This article reports research conducted in a day school and treatment program for autistic children; the nine children who participated in these studies had acquired a basic set of language skills during the earlier stages of their treatment, but needed to develop more elaborate expressive speech. Thus, experiments were designed to evaluate the effectiveness of three language-programming strategies. Experiment 1 examined the effects of a procedure for shaping successively more complex utterances that ultimately included nouns, size/shape and color descriptors and verbs, and assessed generalization across a new teacher, a new classroom, a new set of materials and a new response modality (handwriting). Experiment 2 investigated a procedure for teaching children to answer wh-concept questions (what, why and how), and assessed a strategy for programming response generalization to untrained stimulus materials. Finally, Experiment 3 examined the effects of a program designed to teach children to report on temporally-remote (past) events, using "paragraphic" speech. All three investigations were conducted in children's regular classroom or home environments, during their regularly-scheduled activities, and using materials that are normally available in special education settings and homes. All three language programs were demonstrated to be effective in helping autistic children acquire more complex and sophisticated language skills needed to support their progress toward normalized social participation.

Autistic children typically present a broad range of language problems, including mutism, echolalia, pronoun reversals, perseverative and noncontextual speech and severe delays in receptive and expressive language, to name but a few. The existence of this broad array of language problems implies the need for an equally broad array of language intervention programs.

Requests for reprints may be addressed to Patricia J. Krantz, Princeton Child Development Institute, P.O. Box 2013, Princeton, New Jersey 08540.

Over the past several years, substantial advances have been made in specifying a technology for teaching language to autistic or retarded children who are nonverbal or severely language delayed. A variety of training manuals is now available for teaching verbal imitation and functional speech (cf. Guess, Sailor, & Baer, 1978; Harris, 1976; Lovaas, 1977).

Once children have acquired some functional words and sentences, additional programs are needed to assist them in elaborating their speech. Thus, researchers have reported procedures for teaching children to ask (Twardosz & Baer, 1973) and answer questions (Risley & Wolf, 1967); use plurals (Garcia, Guess, & Byrnes, 1973; Sailor, 1971), prepositions (Sailor & Taman, 1972), and simple and compound sentences (Lutzker & Sherman, 1974: Stevens-Long & Rasmussen, 1974; Stevens-Long, Schwarz, & Bliss, 1976); and sentences with different verb inflections (Clark & Sherman, 1975).

Concomitantly, many investigators have been concerned with programming the generalization of children's newly-acquired language skills across persons and settings (Guess, Keogh, & Sailor, 1978; Handleman, 1979; Sosne, Handleman & Harris, 1979; Warren & Rogers-Warren, 1980) and with developing generalized sets of verbal behaviors, or generative language (Guess, 1969; Guess, Sailor, Rutherford, & Baer, 1968).

Although these analyses have yielded many useful language-intervention strategies, autistic youth who have been in treatment for some period of time, and who have acquired many of the language components mentioned above, often seem to "plateau" at specific skill levels until additional language programs are offered to increase the length, complexity and diversity of their verbal responses. Thus, the research reported in the following includes three experimental analyses of the development of "complex" language: (1) sentences that contain noun, size/shape, color descriptors, and verbs; (2) answers to wh-concept questions (i.e., what, why, how); and (3) "paragraphic" speech about temporally-remote events.

GENERAL METHOD

Participants and Setting

The research reported here was conducted at the Princeton Child Development Institute, a private, nonprofit education and treatment program for autistic children. The nine children who participated in these studies were all enrolled in the Institute's day education program, where they attended school from 9:00 a.m. to 2:30 p.m., five days per week. The eight boys and one girl ranged in age from 5 to 13 years at the time these investigations were conducted. All had been diagnosed autistic by an outside agency, all were severely language delayed, and all displayed an array of severe behavior problems. Eight children lived at home

with their own families, and one resided at the Institute's Teaching-Family group home for autistic youth. Length of time in the program varied from 1 to 4 years, with a mean of 3 years. During each school day, each child had 11 30-minute sessions; at the end of each half-hour, children typically changed teachers, classrooms and activities. This academic schedule is used to promote generalization across persons and settings.

General Procedure

Experiments 1 and 2 were conducted during children's regular academic sessions, while the procedures reported in Experiment 3 were used not only during the school day, but also in children's own homes. Each child enrolled in the program has an individualized motivational system, and the type of system (e.g., contingent edibles, tokens, points on a point card), as well as parameters relevant to density of reinforcement and rate of exchange, are specified in children's case records and are an important dimension of ongoing staff training. In the three experiments reported below, each child's own special motivational system was used to reward desired performances. Thus, the research was conducted in children's usual settings (classrooms and homes), using the activities, motivational systems and naturalistic reinforcers typical of these settings.

EXPERIMENT 1: TEACHING CHILDREN TO USE MULTIPLE DESCRIPTORS

This study examined a procedure for teaching severely language-delayed autistic children to use multiple descriptors in both verbal and written descriptions of objects commonly found at home and school. Previous investigations have assessed the written, receptive and expressive use of adjectives by normal as well as exceptional children. Thus, Maloney and Hopkins (1973) used a point system to modify fourth, fifth and sixth-grade students' use of different adjectives in written stories, and Heward and Eachus (1979) employed a modeling, reinforcement and feedback package to help hearing-impaired and aphasic children use prenominal adjectives in writing sentences. Receptive training of adjectival inflections (e.g., "bigger," "biggest") for three severely-retarded youth was reported by Baer & Guess (1971).

A somewhat larger number of studies deals with children's use of descriptive adjectives in expressive speech. Hart & Risley (1968, 1974) used incidental teaching to increase the use of descriptive adjectives by disadvantaged preschool children. Lahey (1971) increased Head Start children's use of descriptors through modelling without reinforcement, while Martin (1975) used modelling and social reinforcement to help two retarded children use color and size descriptors. A review of the literature reveals no research relevant to autistic children's use of descriptive adjectives, and no examination of their abilities to generalize such

newly-acquired language skills across persons, settings, instructional materials or response modalities. The virtual absence of descriptors from many autistic children's language repertoires, and their characteristic difficulties with generalization, led to the present study.

METHOD

Participants and Setting

Four autistic boys participated in this study. Children 1 and 2 were 7.8 and 8.1 years of age, and achieved respective Mental Age Scores of 2.3 and 5.1 years on the Peabody Picture Vocabulary Test (PPVT), and Vocabulary Age Levels of 2.6 and 3.5 years on the Assessment of Children's Language Comprehension (ACLC). Children 3 and 4 were 12.4 and 13.1 years of age; their Mental Age Scores on the PPVT were 4.2 and 6.2 respectively, and their Vocabulary Age Levels on the ACLC were 5.1 and 5.5 years.

The four children had been in treatment 2 to 4 years, and at the outset of intervention, all had displayed gaze aversion, inappropriate affect, self-stimulatory behaviors, physical and/or verbal aggression and echolalia. Children 1 and 2 had been nonverbal prior to their enrollment in the program. Due to the severity of his behavior problems, Child 3 had been unable to remain at home with his own family, and had lived in a psychiatric institution prior to his enrollment in the Institute's Teaching-Family group home treatment program.

Prior to baseline, an 80-item pretest was used to identify a subset of 13 stimuli that the four children could correctly label with 40 to 100% accuracy. In addition, all four boys achieved 90% accuracy or better on receptive language tasks that required them to identify 7 colors, 4 shapes, 2 sizes and 7 uses of common objects.

Because of differences in skill levels, Children 1 and 2 were paired for daily sessions, and Children 3 and 4 shared separate daily sessions. Thus, each dyad had one 30-minute class per day, five days per week. Sessions were conducted in a small (2.1m $\times$ 3.5m) classroom, and during these sessions, each pair of students was seated facing the teacher. A one-way window and sound system in the classroom permitted observers to stand outside the room when obtaining measures of interobserver agreement with the teacher.

Dependent Measures

The dependent variables in this study were children's use of labels, color descriptors, shape or size descriptors, and verbs. Verbal responses were assessed for all four subjects; subsequently, written responses of Children 3 and 4 were also examined.

Each discrete trial began when the teacher turned to a child, presented an object or a picture of an object, and said, "Tell me about this," or "Tell me about this—write it down." Each verbal or written response was scored for the child's correct use of a label (noun); a color descriptor; a shape or size descriptor; and a verb describing the use or function of the item. Answers were counted as correct only if the child responded without prompts and began his verbal or written response within 3 seconds of receiving the instruction from the teacher. Direct observational data were collected on children's verbal responses; the youth's written answers resulted in permanent products which were later scored by independent observers.

The percentage correct was calculated for each of the following response categories: (1) label (noun) only; (2) color *and* label; (3) shape or size *and* color *and* label; and finally (4) verb *and* shape or size *and* color *and* label. Thus, the response, "This is a yellow pencil" was counted as a correct category 1 and category 2 response, but not as a correct category 3 or category 4 response. The statement, "I can write with a big yellow pencil," was scored as correct in all four categories. This method of organizing the data permitted assessment of the extent to which the training procedure assisted children in acquiring increasingly complex verbal/written language.

During each phase of the study, children remained in each training condition until they achieved 100% accuracy on training items (N = 30) for three consecutive sessions. When this accuracy criterion was achieved, untrained probe items (n = 13) were administered. When a child achieved 90% accuracy or better on three consecutive probes, he entered the next training condition.

Experimental Procedures

The study employed a multiple-baseline design across four categories of language complexity. For Children 1 and 2, these categories were ordered as follows: (1) label; (2) color and label; (3) shape/size and color and label; and (4) verb and shape/size and color and label. To control for effects of order of presentation of descriptors, language categories for Children 3 and 4 were sequenced as follows: (1) label; (2) shape/size and label; (3) color and shape/size and label; and (4) verb and color and shape/size and label.

Baseline. During baseline, five probes were administered to each child. Probes consisted of a set of 13 untrained objects or pictures of objects commonly found in the home and school. Sample probe items included objects such as ruler, pencil, paper, envelope and paper bag, and pictures of wagon, kite, ball, newspaper and radio. Pictures were obtained from the Peabody Articulation Deck. Probe items were randomized prior to each presentation. Children received no feedback on the accuracy of their performances, but were rewarded for answering quickly or sitting quietly during another child's turn.

Teaching Labels. During each training session, two children sat facing the teacher, as in baseline. The teacher began each trial by presenting an object or a picture of an object to a child and instructing, "Tell me about this." If the child provided a correct noun within 3 seconds, and without prompts, he received immediate social praise, plus a token or point, and the teacher began a trial with the other child. In training sessions, as well as during probes, children also earned points or tokens for sitting quietly and for looking at the teacher during another child's turn. After earning predetermined numbers of tokens or points (range = 5-10), children exchanged them for preferred toys, foods or activities.

If a child responded incorrectly, or failed to respond, the teacher modelled the correct response (e.g., "This is a napkin") and then re-prompted with the instruction. "Tell me about this." Following correct, prompted responses, children received behavior-descriptive praise, but did not receive points or tokens. If, after being prompted, a child failed to make a correct response, that trial was terminated, and the teacher presented an object or picture to the other child.

The 30 training items were similar to probe items, and consisted of common objects as well as pictures of objects, again selected from the Peabody Articulation Deck. When a child achieved 100% accuracy on labelling in three consecutive training sessions, the untrained probe items were administered; when 90% accuracy or better was achieved on 3 consecutive probes, the child entered the next condition—Color and Label for Children 1 and 2, and Shape/Size and Label for Children 3 and 4.

Color and Label. Training continued as it had during the preceding condition, except that children were now required to correctly use color descriptions *and* labels (e.g., "This is a white napkin") in order to receive points or tokens. When describing items with more than a single color, children were required to designate the predominant color. If two or more colors were approximately equally represented, children were reinforced for mentioning one or more colors.

Size/Shape and Label. This condition was similar to the "Color and Label" condition described above, except that children were required to use size or shape descriptors instead of color descriptors (e.g., "This is a square napkin"). Size descriptors that had been identified as present in children's receptive vocabularies during pre-baseline assessment were "big" and "little," and shape descriptors that were part of children's receptive repertoires included "round," "square," "rectangle" or "rectangular," and "triangle" or "triangular." When describing objects with several shapes, children were required to describe the primary shape (e.g., a rectangular puzzle with a round piece must be described as rectangular). Responses were counted as correct if a child mentioned *either* the size *or* shape of the item.

As in the "Label" and "Color and Label" conditions, 100% accuracy on size/shape and label over 3 training sessions led to the administration of probes, and

3 consecutive probes with 90% accuracy or better resulted in the initiation of the next training condition.

Size/Shape and Color and Label. Prompting and reinforcement procedures remained the same as earlier, with the exception that children now earned tokens or points for verbal/written productions containing a size or shape descriptor *and* a color descriptor *and* a label (e.g., "This is a square, white napkin"). Requirements for inaugurating probes and for beginning the next experimental condition were identical to those reported previously.

Verb and Size/Shape and Color and Label. In this final condition, children were required to use the previously-trained labels and descriptors, as well as to correctly employ verbs that related to the uses of objects or pictures (e.g., "I look at the big, red book"). Any verb commonly used in conjunction with a particular noun or label was accepted as a correct response, e.g., "I (read/look at/color in) the big, red book." Verbs that did not conform to common social usage were scored as incorrect, e.g., "I play with the big, red book," or "I read the little, yellow puzzle."

Reliability

Interobserver agreement on children's verbal responses was obtained during all 5 baseline probes for each of the 4 children; on 9 of 16 posttraining probes for Children 1 and 2; on 5 of 11 posttraining probes for Children 3 and 4; and on one-month follow-up probes for Children 3 and 4. An agreement was scored for a given trial only if the observers agreed on label *and* size/shape *and* color *and* verb. Interobserver agreement was calculated as follows: number of agreements/number of agreements + disagreements $\times$ 100. Interobserver agreement was 100% on all 14 reliability checks for Children 1 and 2, and 100% on all 11 reliability checks for Children 3 and 4.

Interobserver agreement on the written responses of Children 3 and 4 was obtained by providing children's written work to independent observers, who used the same data sheets and scoring procedures that were employed in collecting data on children's verbal productions. Reliability checks were obtained for all 4 baseline probes, and for 5 of the 10 posttraining probes. Again, interobserver agreement was 100% for both children, and across all conditions.

RESULTS

Figures 1 and 2 display the percent correct responses to probes for each of the four children, at each level of language complexity, and across the total number of (training and probe) days of the study. As mentioned previously, probes were administered during baseline, and when children met the 100%

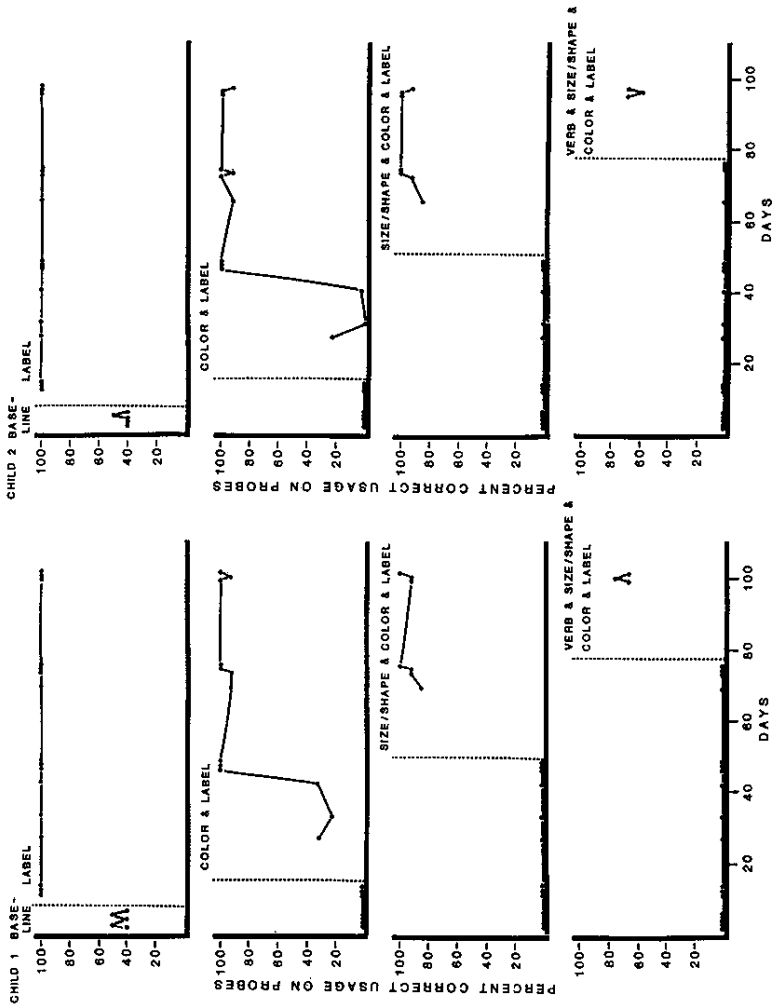

FIGURE 1. Percent correct usage for each level of oral language complexity on 13 untrained probe items presented to Children 1 and 2 before and after teaching.

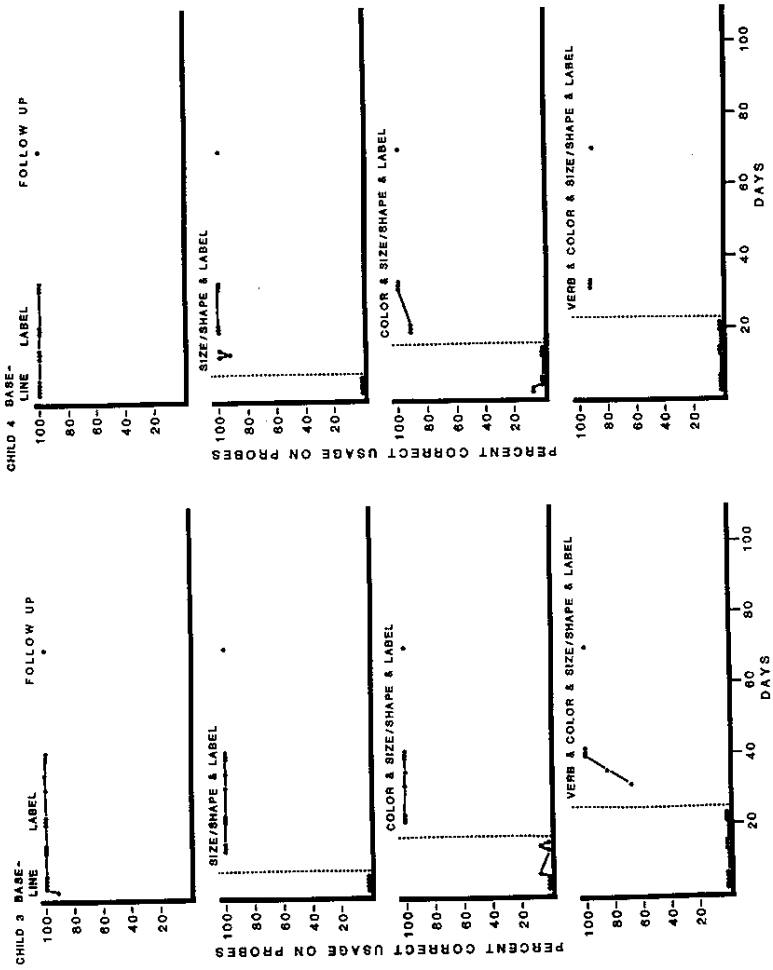

FIGURE 2. Percent correct usage for each level of oral language complexity on 13 untrained probe items presented to Children 3 and 4 before and after teaching, and at one-month follow up.

accuracy criterion on training items for three consecutive sessions. During baseline probes, Children 1 and 2 provided correct labels for 40 to 50% of the items presented, while Children 3 and 4 correctly labelled 92 to 100% of pre-training probe items. However, neither Child 1 nor Child 2 ever used a correct color descriptor, size/shape descriptor or verb until these were trained, and Children 3 and 4 provided only a few descriptors (0 to 8% correct on "Color and Size/Shape and Label") prior to training.

Posttraining probes indicated that all of the children quickly achieved criterion of 90% correct or better for three consecutive probes in each condition. Once criterion was achieved, children's performances were maintained throughout the study, and continued to maintain at a one-month follow-up probe (Day 70) for Children 3 and 4. Follow-up probes were not obtained for Children 1 and 2, because they did not attend the summer continuation of their regular academic programs.

Generalization Probes

Pre- and posttraining generalization probes were employed in order to assess transfer of treatment effects to another teacher, another classroom, another set of materials (a new set of untrained probe items) and another response modality (written, rather than verbal responses to probe items). Data collection procedures were as previously described; interobserver agreement was obtained on each of these eight probes, and remained at 100% for each of the four children.

In order to assess generalization across teachers, a new teacher was asked to administer the standard probe items. Table 1 indicates that the children used only labels during the pretraining probe. On the posttraining probe, Children 1 and 2 achieved 67 and 59% respectively, on "Verb and Size/Shape and Color and Label," but otherwise, all previously-trained responses remained at 100% correct on probe items for all four children.

To examine generalization across classrooms, the session was moved to a different room, but standard probe items were administered by the same teacher. On this pretest, the children used only nouns; in the posttraining probe, Children 1 and 2 again showed some decrements in performance on the most complex language category (67% correct), but all other responses remained at 100% correct for all four children.

Generalization across materials was investigated by providing a different set of 13 untrained probe items that, as previously, included both objects (e.g., shoebox, bottle, mirror, tissue) and pictures of objects (e.g., magazine, comb, toothbrush and top) that were obtained from the Peabody Articulation Deck. On this pretraining probe, the four youth continued to respond only with labels, while on the posttraining probe they achieved 100% correct on the first three levels of language complexity and 59 to 85% correct on the fourth level of language complexity. Thus, although they showed some decrements in providing

TABLE 1
Pre-and Post-Training Probes on Generalization to New Teacher, New Classroom, New Materials and New Response Modality (Writing)

Level of Language Complexity	New Teacher				New Classroom				New Materials				New Response Modality (Writing)			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
Pre-Training																
Label	40	50	100	100	10	40	100	100	67	77	100	100	–	–	100	100
Color and label	0	0	0	0	0	0	0	0	0	0	0	0	–	–	0	0
Size/shape and color and label	0	0	0	0	0	0	0	0	0	0	0	0	–	–	0	0
Verb and size/shape and color and label	0	0	0	0	0	0	0	0	0	0	0	0	–	–	0	0
Post-Training																
Label	100	100	100	100	100	100	100	100	100	100	100	100	–	–	100	100
Color and label	100	100	100	100	100	100	100	100	100	100	100	100	–	–	0	0
Size/shape and color and label	100	100	100	100	100	100	100	100	100	100	100	100	–	–	0	0
Verb and size/shape and color and label	67	59	100	100	67	67	100	100	59	67	85	85	–	–	0	0

verb and size/shape and color and label, overall, their performances maintained after training.

Finally, the original set of probe items (objects and pictures of objects) was administered to Children 3 and 4, but the children were asked to use a new response modality—written, rather than verbal answers. The teacher used the instruction, "Tell me about this—write it down." This generalization probe could not be administered to Children 1 and 2 because, at the time of the study, they had not yet acquired handwriting skills.

Table 1 shows that Children 3 and 4 were able to provide correct written labels for stimulus pictures and objects, but were unable to demonstrate any higher levels of language complexity in their written responses, either before or after training on their verbal responses. This finding led to a replication of the previously-reported research, using a different response modality (handwriting).

Replication Across Response Modalities

In order to assure that the failure of Children 3 and 4 to generalize from verbal to written responses was not attributable to their spelling and writing skills, these skills were formally assessed. Both youths were able to correctly print and spell common nouns and verbs, color names and shape/size words with 90 to 100% accuracy.

During baseline and posttraining probes, the teacher presented an object or picture of an object, and said "Tell me about this—write it down." These probe items were similar to those used for verbal response training, and were selected on the basis of preliminary assessment data indicating that the children could correctly spell and write the required words. Children received no feedback on the accuracy of their written responses to probe items, but did receive points on their point cards for working quickly and quietly.

During training, the teacher continued to present objects or pictures and to give the instruction, "Tell me about this—write it down." In training sessions, however, the teacher checked each written response immediately after it was completed. Correct, unprompted responses were rewarded with behavior-specific praise and points. If a youth produced an incorrect written response, however, the teacher modelled (wrote) the correct answer and asked the child to correct his paper. Corrected responses received teacher praise, but were not followed by delivery of points.

Figure 3 presents children's performance on written, rather than verbal responses. A small change in data presentation was necessary; because children required more time to write their responses than to provide verbal responses, probe data are presented in blocks of 10 trials. Otherwise, all procedures remained the same as previously described. Children 3 and 4 used virtually no adjectives or verbs prior to training, but subsequent to training, the children achieved high levels of accuracy on probes, and were able to provide written responses that included correct verbs, size/shape and color adjectives, and labels.

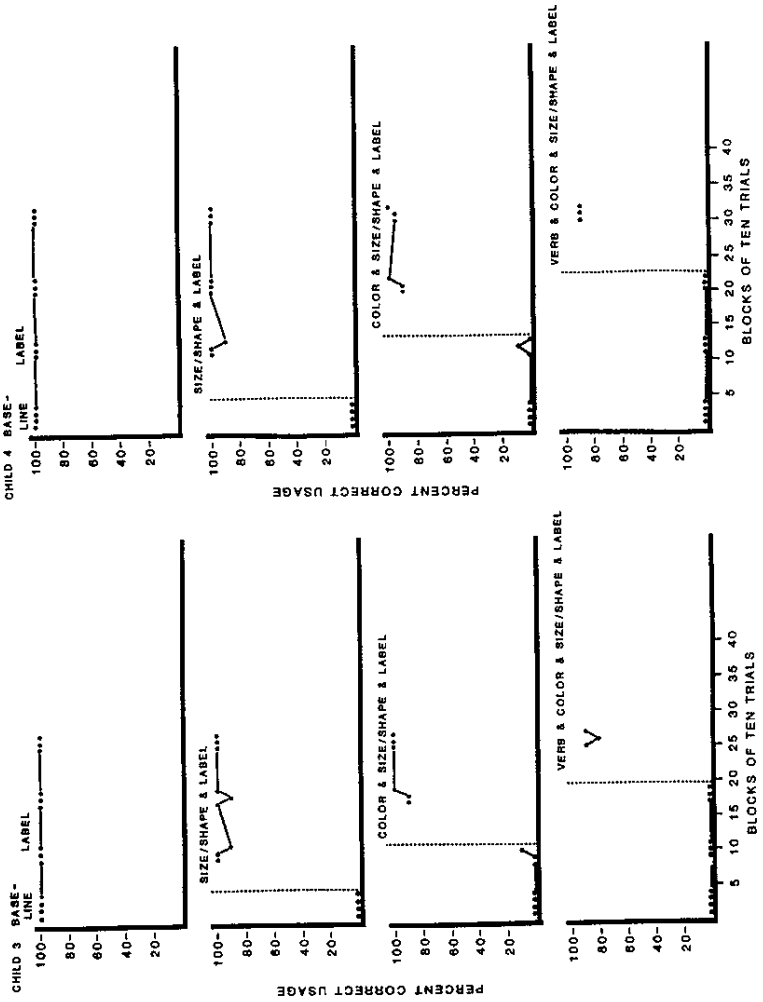

FIGURE 3. Percent correct usage for each level of written language complexity in 10 untrained probe items presented to Children 3 and 4 before and after teaching.

DISCUSSION

The experimental procedures described above were successful in increasing the complexity of four autistic children's expressive speech. Initially, all four children, like many other autistic youth, responded to stimulus materials with nouns only. Subsequent to training, however, the children were able to use appropriate nouns, verbs, shape/size, and color adjectives. Since the training procedure was essentially a shaping procedure, children maintained each previously-taught language component, while adding new components to their verbal repertoires.

Data collected immediately after the completion of expressive language training indicated that all of the children generalized their new skills across a new teacher, a new classroom and a new set of stimulus materials. However, Children 3 and 4, who were known to have adequate handwriting and spelling skills, were unable to generalize from oral to written language. This finding led to a replication of the training procedure for written, rather than spoken responses, with the result that both children achieved a level of written language complexity similar to that demonstrated earlier for expressive speech.

Anecdotal reports from teachers and parents suggested that the children's new language skills were maintained, and that they generalized to other persons and settings. Thus, one child's parents reported that he had begun to use more complex speech at home, e.g., "I want to drink from the white glass" rather than "juice;" "I want to bounce the ball" rather than "I want ball."

The study suggests that, although some autistic children's receptive vocabularies may include verbs and adjectives, these may not appear in children's expressive speech in the absence of a special prompting and reinforcement procedure. Prior to training, children appeared to respond with "minimum effort," primarily using only nouns to manipulate their environment. When reinforcement was contingent upon increasingly sophisticated verbal output, however, the children's verbal repertoires increased in complexity. This was regarded as a step toward more normalized and "conversational" speech.

EXPERIMENT 2: TEACHING ANSWERS TO WH-CONCEPTS

Brown (1968) conducted a naturalistic study of the ability of three preschool children to answer wh-questions, and reported that these skills were acquired between the ages of 2.5 and 5 years. He wrote, "The knowledge represented by transformational grammar is not, of course, explicitly taught by parents, but must somehow be derived by the child from linguistic data." Autistic children, however, usually do not learn language through the same patterns of casual observation, verbal experimentation, and practice displayed by normal youngsters, and require formal instruction in order to acquire speech skills.

Gallagher and Darnton (1978) and Twardosz and Baer (1973) reported pro-

cedures for teaching language-delayed children to ask wh-questions, and Wilcox and Leonard (1978) used pictures as visual prompts for question asking. Lovaas (1977) discussed prompting and reinforcement procedures used to teach autistic children to ask and answer social questions, and presented data on children's acquisition of question asking (e.g., "What's your name?" or "What are you going to eat for dinner?").

Responding to wh-questions is a basic communication skill, and one that greatly facilitates children's participation in normal social interaction. However, there have, as yet, been few investigations of procedures for teaching autistic children to answer specific classes of wh-questions.

METHOD

Participants and Setting

The three autistic boys who participated in this research had previously acquired certain prerequisite skills, such as following simple directions, and using simple sentences that included nouns, verbs, descriptive adjectives and pronouns, e.g., "I want the big ball" or "My coat is blue."

Child 5 was 10 years of age, and had been enrolled in the program for 2½ years prior to the beginning of the study. At admission, he achieved a Mental Age Score of 6.8 on the Peabody Picture Vocabulary Test (PPVT), but engaged in virtually no spontaneous speech. He was a hyperlexic child who had learned reading recognition skills at age 2.5, but by age 7 reading comprehension (and other language functions) continued to be very delayed.

Child 6, age 9, had been nonverbal when he entered the program three years earlier, and had been unable to achieve a basal score of six consecutive correct responses on the PPVT. His presenting problems included phobic responses to daily activities such as bathing and hair brushing; attempts to preserve "sameness" in his environment (e.g., demands to wear the same article of clothing every day); and self-stimulatory behaviors such as twirling.

Child 7, age 5, was also unable to achieve a basal score on the PPVT administered at the time of his enrollment a year earlier. At the outset of intervention, he engaged in self-stimulatory handwaving and high levels of crying.

Throughout the period of study, each child had a daily 1:1 session with the teacher. Sessions were 30 minutes in length, and were conducted at the same times of day, five days per week, in a typical classroom environment. The child was seated directly in front of the teacher, and the teacher presented stimulus materials in a discrete-trial format.

Overview

Adverbial interrogative words (when, where, why, how) and nominal interrogative words (who, whom, whose, what, which) are here referred to as "wh-concepts." Each of the wh-concepts named above can be divided into many "subconcepts," and it is important to identify the most commonly-used subconcepts, and to train these individually, in order to help children learn to respond to a variety of usages of wh-words. For this reason, three wh-concepts on which the target children were deficient were divided into subconcepts that reflect common social usage, as described in the following paragraphs.

What. This concept was divided into four components, as follows: (1) What as an object or noun, e.g., "What is in this picture?"; (2) What as an infinitive verb phrase, e.g., "What is the girl doing?"; (3) What as a predicate, e.g., "What is the boy eating?"; and (4) What as "which" e.g., "What game are the men playing?".

How. This interrogative is the only one of the wh-forms that does not begin with the letters "wh" (see, for example, Quirk & Greenbaum, 1973). The "how" concept was divided into five subconcepts, including: (1) How relevant to facial expression, e.g., "How does the girl look?"; (2) How as in "How many...?"; (3) How as in "How often...?"; (4) How related to action, e.g., "How will the men get to the top of the mountain?"; and (5) How meaning "by what means", e.g., "How do the people keep warm?".

Why. This interrogative word was broken into four commonly-used categories: (1) Why relevant to cause/effect, e.g., "Why is the man wearing a coat?"; (2) Why in reference to affect, e.g., "Why is the boy laughing?"; (3) Why related to potential action, e.g., "Why is the family in the car?"; and (4) Why questions about natural occurrences, e.g., "Why do birds have wings?".

These subcategories do not reflect all of the separate usages of these wh-forms, nor do they always reflect grammatical classifications, e.g., "why" referring to affect is not grammatically different from "why" referring to cause and effect. However, for purposes of programming instruction for autistic children, it appears important to define these as separate classes of stimuli to be trained, because many children, after learning to answer "why" questions about emotional affect, would continue to be unable to answer "why" questions about people's performance of the tasks of everyday living (e.g., wearing coats, brushing teeth, stopping at a red light). Thus, "subconcepts" were identified on the basis that they called for verbal responses commonly used in conversational speech.

Dependent Measures

The dependent variables in this study were children's responses to "what," "why" and "how" probe questions about magazine pictures. Children's answers to wh-questions were counted as correct only if they were unprompted, and if they occurred within 10 seconds after a question was posed.

During each 1:1 session in which probe items were asked, a tape recorder was used to obtain children's responses to the wh-questions. Then, independent observers were given the tapes and the complete set of probe pictures and questions, and were asked to transcribe verbatim the children's responses to probe questions. Subsequently, observers scored each child response on the transcript as correct or incorrect.

Since any single probe question might produce several correct responses, observers scored children's answers as correct if they judged them to be "reasonable" and "socially appropriate." For example, a question such as "Why does the boy go to the dentist?" might produce answers such as, "Because his teeth hurt," "Because his mother took him to the dentist," or "Because he will have his teeth cleaned," and all of these would be scored as correct. Examples of answers that would not be scored as correct include "Because he is little," "The man," or "The coat is white."

Single-Concept Probes. When the teacher had completed training on all subconcepts that were specified for a particular concept, she administered a single-concept probe. These probes consisted of 20 untrained wh-questions over 10 new magazine pictures (two questions for each picture). The 20 items for each concept were distributed equally across subconcepts. For example, the 20 items for the wh-form "what" included 5 items concerning "what" as a noun/object; 5 items on "what" as an infinitive verb phrase; 5 items on "what" as a predicate; and 5 items on "what" meaning "which." None of the pictures or questions that comprised the single-concept probes was ever included in training sessions or presented for any purpose other than to obtain single-concept probe data. Thus, children's performance on these probe items provided a measure of response generalization from trained to untrained stimulus materials.

In most cases, the child was required to achieve 80% correct or better on at least two out of three single-concept probes before the teacher administered an all-concept probe and began training on the next wh-concept. If a child did not achieve 80% correct on a single-concept probe, two or more training sessions were conducted before the probe was re-administered. In some cases, the teacher continued to train and retest on the single-concept probes even though a child had achieved 80% correct on two of three probes. This happened because the teacher elected to provide additional training designed to assist the child in answering questions with complete sentences.

All-Concept Probes. Like the single-concept probes, these probes consisted of 10 magazine pictures, each accompanied by two questions, for each wh-concept. Thus, the all-concept probes consisted of 60 questions, 20 each for "what," "why" and "how." As in the case of single-concept probes, items were equally distributed across subconcepts. Items included in the all-concept probes were never trained; thus, these probes also provided a measure of generalization across materials. All-concept probes were administered at least three times during baseline, and before beginning each new teaching condition.

Experimental Procedure

A multiple-baseline design across wh-concepts was used with each of the three children. Child 5 received instruction on "what," then on "why," and finally, on "how." Children 6 and 7 were trained first on "what," and then on "how," and lastly, on "why." This variation in the sequencing of wh-forms was used to control for possible effects of order of presentation.

Baseline. During baseline, all-concept probes were administered three or more times. The 60 items comprising the all-concept probe (20 for each wh-concept) were untrained. When presenting a picture and asking a wh-question, the teacher delivered points contingent upon the child's appropriate classroom demeanor, e.g., sitting quietly in his chair, visually attending to the teacher and the stimulus materials, and keeping his hands in his lap. However, children received no feedback on the accuracy or correctness of their verbal responses. At the end of each session, children exchanged their points for preferred activities, foods, or play materials.

Teaching. The instructional materials used were a variety of magazine pictures selected by the teacher; none of these pictures was the same as any of the pictures included in the probes. Teaching began on a specific subconcept of the first wh-concept. In providing instruction, the teacher presented a magazine picture (e.g., a girl drinking milk) and asked a wh-question such as "*What* is the girl drinking?". Initially, two types of prompts were used—the teacher placed verbal emphasis on the wh-form and, if possible, pointed to an object in the picture (e.g., milk) that represented a correct response. As children became more proficient in using a subconcept, these prompts were faded.

In addition, the teacher provided verbal prompts that assisted the child in identifying a general response category that was being requested. For example:

Teacher: *How* does the boy look?
Child: I don't know.
Teacher: Does he look happy?
Child: No.
Teacher: Does he look sad?

Child: No.
Teacher: *How* does the boy look? Does he look happy, sad or angry?
Child: The boy looks angry.
Teacher: (Delivering a point) Good, you told me *how* the boy looks.

These verbal prompts were also faded as children acquired more skill in answering questions about a particular subconcept.

During training, the teacher also prompted and reinforced the use of complete sentences in answering wh-questions. Initially, children responded to many of the questions with only words or phrases, but this was unacceptable because it did not permit the teacher to ascertain whether they were responding to the particular wh-form being taught, or to some other wh-form. By way of illustration, the teacher might ask, "How do the men cross the water?" and if the child replied "Boat," it could not be determined whether he was responding to the wh-form "how," or to some other wh-form such as "What is this?". Thus, the teacher prompted by instructing, "Say the whole thing," or "Say it in a sentence," and if these levels of prompts were ineffective, she modelled the sentence and asked the child to imitate. Prompts for the production of complete sentences were also faded as training proceeded, so that as children approached criterion performances, they received points only when they gave unprompted answers in complete sentences. During training sessions, children were permitted to exchange points for back-up reinforcers within, as well as at the end of sessions.

When training data collected by the teacher indicated that a child had achieved mastery of a subconcept, a subconcept probe was delivered. These probes consisted of 20 untrained questions about 10 untrained magazine pictures. All 20 items related to the particular subconcept that had just been trained, e.g., "why" as cause/effect. If the child achieved 80% accuracy or better on the subconcept probe, training began on a new subconcept. If he failed to achieve 80% accuracy, training on the subconcept was resumed for two or more sessions before the subconcept probe was again administered. The procedure used during the delivery of subconcept probes was similar to that used for single-concept probes and all-concept probes, i.e., the teacher gave children points contingent upon their appropriate classroom behavior, but did not provide feedback about the accuracy of their responses.

Subconcept probes were considered training data, and were used as feedback to the teacher concerning the child's progress. These probes also provided an informal measure of children's ability to generalize to a new set of stimulus materials. Interobserver agreement was not obtained for subconcept probes.

Teaching continued in the manner detailed above until a child had achieved 80% correct or better on all of the subconcept probes associated with a particular wh-concept. Then the teacher administered the single-concept probes and all-concept probes as described earlier. Figure 4 displays a summary of the sequence of measurement and intervention activities.

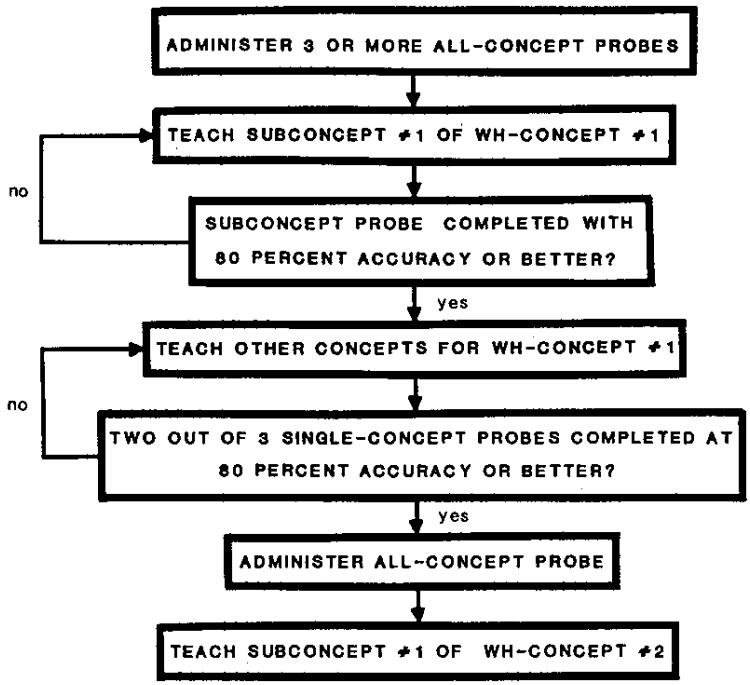

FIGURE 4. Summary of measurement and intervention activities in the wh-concepts program.

Reliability

During each single-concept or all-concept probe session, an audio tape of the child's responses to wh-questions was obtained. This tape was given to an observer, together with the relevant magazine pictures and questions. The observer first made a transcript of all child responses on the tape, and subsequently, scored each child response as correct or incorrect. Later, the same materials were given to a second naive observer, who performed the same operations. The formula used to calculate agreement was: total number of agreements/total number of agreements + disagreements × 100.

Interobserver agreement on single-concept probes was obtained for 9 of Child 5's sessions; 10 of child 6's sessions; and 11 of Child 7's sessions. Agreement ranged from 80 to 100% for Child 5, with a mean of 87; from 75 to 100% for Child 6, with a mean of 84; and from 70 to 100% for Child 7, with a mean of 84.

Interobserver agreement on all-concept probes was obtained for all 7 probes for each of the three children. Mean percent interobserver agreement for Child 5 was 86 (range = 80 to 100%); the mean for Child 6 was 88 (range = 80 to 100%); and the mean for Child 7 was 88 (range = 80 to 100%).

RESULTS

Figures 5–7 display children's accuracy in answering wh-questions on all-concept probes, before and after teaching. It may be noted that during baseline, all three children scored higher on "what" than on "why" or "how," perhaps because more "what" questions are used in their daily academic curricula (e.g., "What letter/number/color/word/shape is this?" or "What do you want to buy with your tokens?").

Child 5 (Figure 5) achieved baseline means of 67% for "what," 26% for "why," and 34% for "how." After training, however, his means increased to 95, 77 and 80% correct for "what," "why" and "how," respectively.

Figures 6 and 7 indicate that these results were replicated for the other two boys. After learning the wh-form "what," Child 6 may have generalized slightly to the two other wh-concepts, but this was not observed for Children 5 and 7. It is especially noteworthy that once a wh-concept was taught, children's skills in answering questions about that concept were maintained.

Examination of the children's performance on single-concept probes, as well as on all-concept probes, indicates that the three boys succeeded in generalizing their newly-acquired question-answering skills across both sets of untrained stimulus materials.

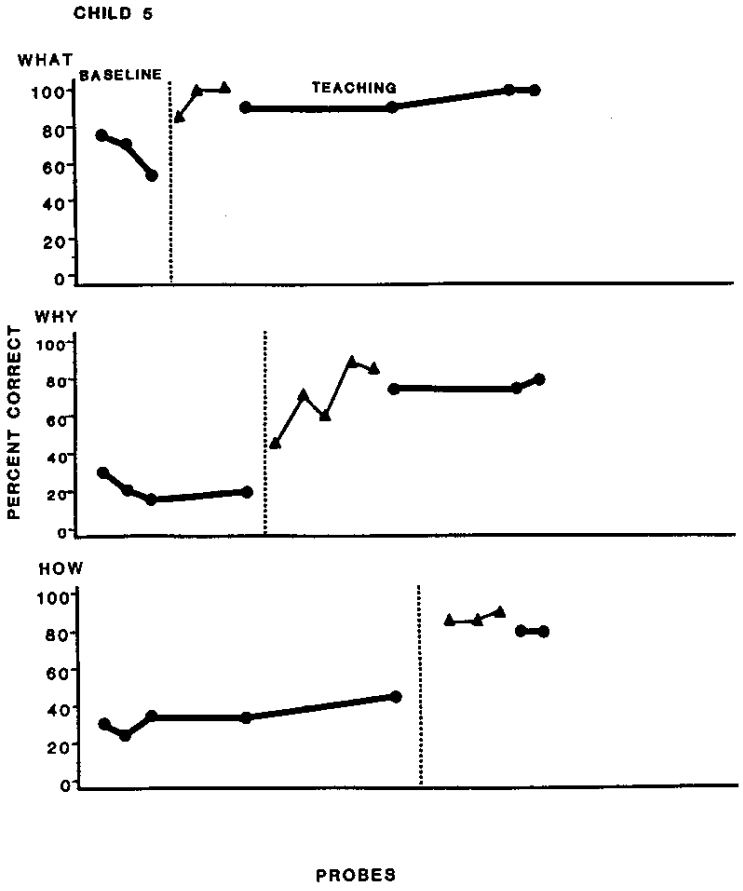

FIGURE 5. Child 5's percent correct answers to untrained "what", "why" and "how" probe questions, before and after instruction. Sixty-item all-concept probes are represented by circles, and 20-item single-concept probes are represented by triangles.

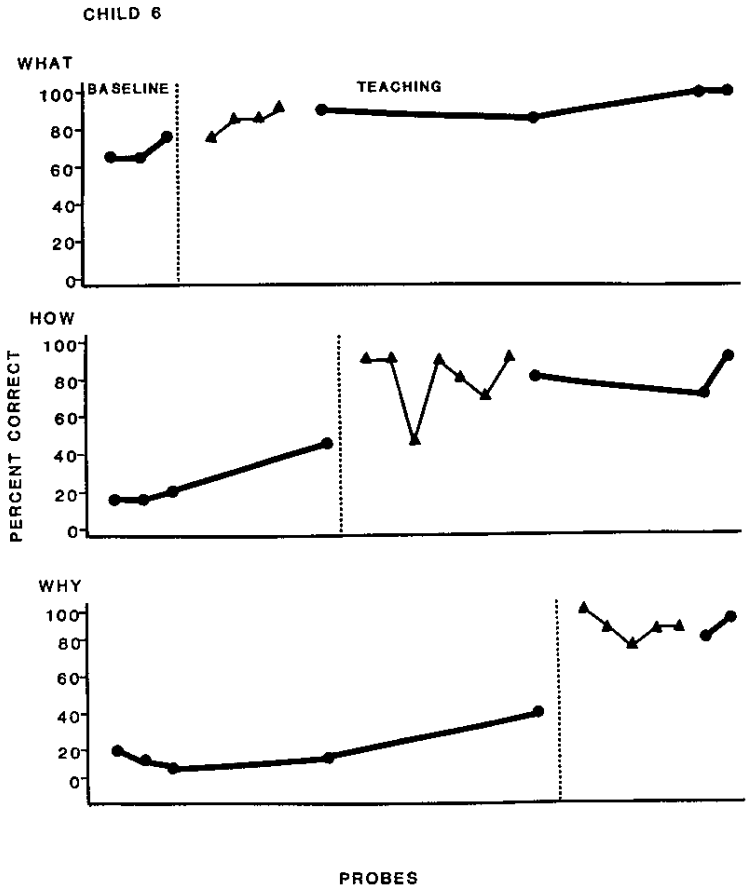

FIGURE 6. Child 6's percent correct answers to untrained "what", "how" and "why" probe questions, before and after instruction. Sixty-item all-concept probes are represented by circles, and 20-item single-concept probes are represented by triangles.

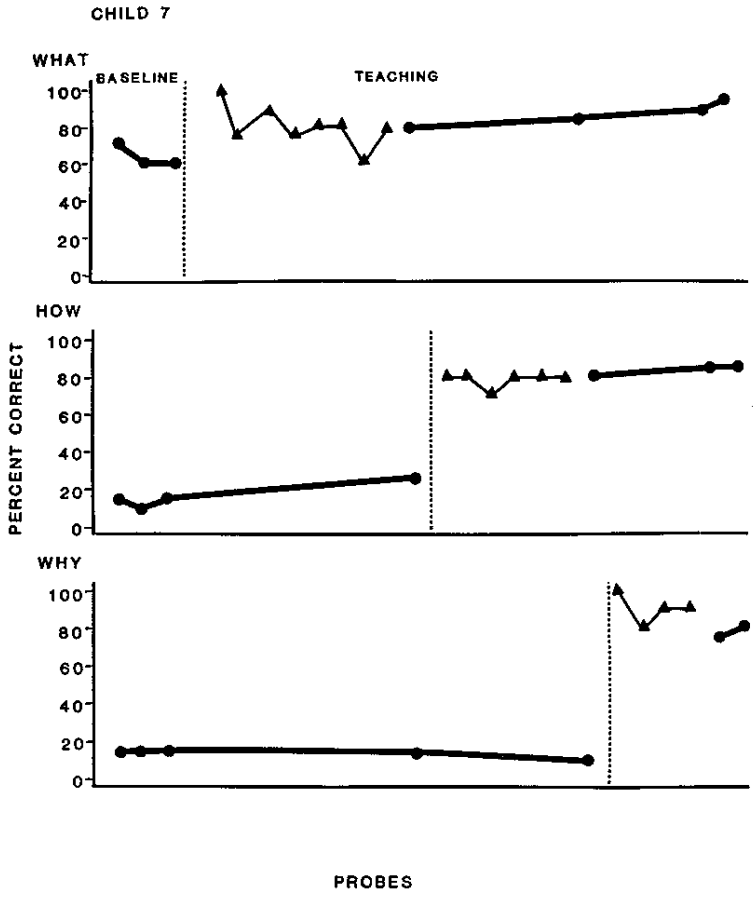

FIGURE 7. Child 7's percent correct answers to untrained "what", "how" and "why" probe questions, before and after instruction. Sixty-item all-concept probes are represented by circles, and 20-item single-concept probes are represented by triangles.

DISCUSSION

Preliminary assessment and informal observation indicated that the three target children initially experienced little success in answering wh-questions. If shown a picture and asked a question such as, "Why are the men wearing swim suits?", the boys either appeared to "guess" at the answer (e.g., "ocean," "blue water," "red suit", etc.) or, pursuant to their earlier training (cf. Schreibman & Carr, 1978) they replied, "I don't know."

Children's response latencies also seemed to be indicators of the initial difficulties they had with wh-forms; a relatively long permissible latency was specified for probe sessions (i.e., children had to begin their answers within 10 seconds from the time the teacher asked the questions, in order for answers to be counted as correct). Informal observation suggested that children's response latencies decreased dramatically after training.

Another favorable result of the teaching program was its impact on children's use of complete sentences. Transcripts of the boys' responses to all-concept probes were used to recover data on the percent of answers made in sentences (where "sentence" was defined as a statement that contained both a subject and a verb). Twenty-seven percent of Child 5's answers were provided in sentences before training, while 80% of his answers were given in sentences after training (mean interobserver agreement on presence/absence of sentences was 80%). Similar effects were observed for Children 6 and 7.

As mentioned earlier, teaching children to respond in complete sentences was necessary, so that it could be ascertained whether they were responding to the wh-form contained in the question, or to some other wh-form. During the course of the study, however, it appeared that the teaching procedure was functional not only because it served to increase the length of children's verbal productions, but also because it contributed to the development of response chains, in which the question prompted an initial part of the answer, and the first part of the answer prompted the remainder. For example, the teacher might ask, "Why is the horse drinking?" and the child might respond, "The horse is drinking because he is thirsty." Or a question such as, "How did the girl get her shirt dirty?" might be answered, "The girl got her shirt dirty because she spilled her ice cream." Most of the wh-forms of the English language are conducive to this formula, because they ask for identification of the subject, object, complement or an adverbial of a sentence (Quirk & Greenbaum, 1973). For example, a picture may appear to have the theme, "The family eats breakfast every morning," and this may lead to several questions, including: "Who eats breakfast every morning?", "What does the family eat every morning?", and "When does the family eat breakfast?". Each of these questions contains phrases that serve as prompts for an initial portion of an answer. Teaching children to make use of such prompts may contribute to the maintenance of newly-acquired skills.

The data obtained from single-concept probes are interesting because they

further document response maintenance on untrained stimuli. At face value, this may appear curious, since children were never reinforced for responses to probe items. However, it may be that the teaching procedure served to shape those child responses that underlie all of their attempts to answer wh-questions. Thus, although children were nonreinforced for answering probe items, the critical set of skills necessary for responding to wh-forms continued to be reinforced during teaching sessions (cf. Siegel & Spradlin, 1978).

An essential ingredient of this intervention procedure is a technology for producing response generalization; this technology has been identified by Stokes & Baer (1977) as "training sufficient exemplars." The procedure described in previous sections provides many opportunities for children to generalize to new stimulus materials. The most important of these opportunities occurs during training, when the teacher selects an unspecified (but usually large) number of pictures to use as training materials. When children appear to have mastered these, they are presented with a new set of materials (a subconcept probe). If criterion is not met on this probe, the teacher may elect to identify additional training pictures and teach children to respond to them, prior to the next probe. If criteria on subconcept probes are successively met, the child soon encounters a new set of stimulus materials, the single-concept probes, and training sessions continue until the child can successfully generalize to these. When criteria are met on single-concept probes, additional sets of untrained materials (all-concept probes) are presented. Informal observation and anecdotal evidence suggest that this generalization strategy is highly effective in enabling children to answer a broad range of wh-questions.

The teaching procedure described above appeared to be economical in terms of time; Children 5 and 6 completed the program in 2 months (36 teaching sessions) and Child 7 completed the program in 3 months (54 teaching sessions). During the past 3 years, a dozen children have participated in this program, with results similar to those described above.

EXPERIMENT 3: TEACHING CHILDREN TO ANSWER QUESTIONS ABOUT TEMPORALLY-REMOTE EVENTS, USING "PARAGRAPHIC" SPEECH

A study by Frank, Allen, Stein, & Myers (1976) compared the verbal interactions of autistic children and their mothers with normal children and their normal mothers, and normal children and their schizophrenic mothers. Although the quantity and complexity of language used by the mothers of autistic children was equal to, or greater than, the language used by other mothers, it was noted that the questions asked by the mothers of autistic children differed from questions asked by other mothers, in that they were predominantly in the present tense. This may have been due to the mothers' perceptions of their children's difficulties in responding to questions about temporally-remote events.

Autistic children's characteristic lack of skills in answering questions about past and future happenings has, as yet, received only minimal research attention. Boucher and Warrington (1976) reported some similarities between the memory deficits of autistic children and the amnesic syndromes of adulthood. Hermelin (1971) found that autistic children achieved equal success in recalling random series of words and series of words that were categorically related (e.g., kitchen utensils), and in recalling ungrammatical and anomalous vs. meaningful grammatical sentences. In sum, the autistic children, unlike their normal and retarded peers, experienced equal success in recalling "sense," or meaningful information, and "nonsense," and this finding applied to the recall of both auditory and visual materials.

Later research (Hermelin, 1976) indicated that when recalling a series of three digits, the autistic children studied invariably used spatial configuration cues rather than temporal cues, i.e., cues related to the order in which stimuli were displayed. It was hypothesized that temporal-ordering skills are related to the use of language to extract rules and to code and store information, and that autistic children experience difficulty with temporal ordering because of their language deficits (Fay & Schuler, 1980).

As yet, there have been few attempts to design and evaluate intervention strategies that may assist autistic children in temporally structuring information. The present study assessed the effects of a verbal rehearsal procedure in helping children answer questions about temporally-remote (past) events, using "paragraphic" speech.

METHOD

Participants and Setting

At the time of this study, Children 8 and 9 were in their first year in the school program, and both lived at home with their own parents.

Child 8, a five-year-old boy, achieved a Mental Age Score of 2.2 on the Peabody Picture Vocabulary Test (PPVT) administered 8 months prior to the study, and a vocabulary age level of less than 3.0 on the Assessment of Children's Language Comprehension (ACLC) administered 5 months prior to the study. His target behaviors included echolalia and perseverative speech; disruptive behaviors associated with changes in daily routines; oppositional responses to adult instructions; tantrums; and self-injurious behavior. His school program emphasized development of greeting skills and descriptive language; simple direction following; and beginning math, handwriting and reading comprehension skills.

At the time of her enrollment in the program and five months prior to the beginning of the study, Child 9, a five-year-old girl, was unable to achieve a

basal score on the PPVT. Two months before the study, she achieved an ACLC vocabulary age level of less that 3.0 years. Her presenting problems included self-stimulatory behaviors such as putting her fingers in her ears and making vocal noise; tantrums; and frequent crying and screaming associated with changes in her usual schedule. Her instructional program was similar to that of Child 8, emphasizing direction following and basic language and academic skills.

Data collection and intervention procedures were conducted in children's classrooms during the regular school day, and in children's own homes after school.

Dependent Measures

The target responses in this study were children's answers to questions about activities that they had observed or participated in earlier that day, or on the previous day.

Each day, teachers wrote questions about children's activities during the school day and provided written answers to these questions; similarly, parents wrote questions about the youngsters' activities at home on the previous afternoon/evening and provided correct answers to these questions. Questions about school activities were sent home in children's lunchboxes, and were asked by

TABLE 2
Sample Data Sheet Showing some Typical Questions and Answers Written at School, Enabling Parents to Ask Questions about Temporally-Remote School Events.

CHILD'S NAME : Carol DATE : 5/26/81

QUESTION AND ANSWER	RESPONSE FROM CHILD	SCORING
Q-1 WHAT DID YOU HAVE FOR LUNCH? A-1 I HAD BOYSENBERRY YOGURT. IT IS BLUE.	boysenberry	CORRECT WORD
Q-2 WHO MET YOU AT THE BUS? A-2 MR. BROWN MET ME AT THE BUS. TOMMY WAS THERE TOO.	no response	INCORRECT
Q-3 WHAT DID YOU DO IN ART CLASS? A-3 I PAINTED WITH RED PAINT. I LIKE TO PAINT.	You painted.	CORRECT SENTENCE

TABLE 3
Sample Data Sheet Showing some Typical Questions and Answers Written at Home, Enabling the Teacher to Ask Questions about Temporally-Remote Home Events.

CHILD'S NAME : Billy DATE: 6/8/81

QUESTION AND ANSWER	RESPONSE FROM CHILD	SCORING
Q-1 WHAT DID YOU DO WHEN YOU CAME BACK FROM SCHOOL? A-1 I ANSWERED THE QUESTIONS. I ATE CEREAL.	school	INCORRECT
Q-2 WHAT DID YOU DO OUTSIDE WITH YOUR BROTHERS? A-2 WE PLAYED WITH THE SPRINKLER. I WORE MY BATHING SUIT.	I jumped over the sprinkler.	CORRECT SENTENCE
Q-3 WHAT TV SHOW DID YOU WATCH IN THE EVENING? A-3 I WATCHED "SHA NA NA" IT WAS FUNNY.	I watched "Sha Na Na". It was funny.	CORRECT PARAGRAPH

parents. Questions about home activities were sent to school, and were asked by teachers.

Children's initial, unprompted responses to each question were recorded verbatim by the questioner (a teacher or parent). Each child response was later scored as correct/incorrect, and as a word or phrase, sentence or paragraph. Tables 2 and 3 provide examples of questions and answers written by teacher and parents, and of the method of recording child responses.

A child response was scored as a sentence if it contained a subject and a verb; an answer was scored as a paragraph if it contained two or more sentences; and a response that did not conform to the definition of a sentence was scored as a word or phrase.

Since many questions had the potential for generating more than a single correct answer, child responses were scored as correct if they matched the content of the written answers provided by teacher or parents, or if they were judged to be contextual and socially appropriate. Only answers that were correct in terms of content were scored for length. Thus, an inaccurate or noncontextual sentence would not be counted as correct, even though it possessed both subject and verb.

Finally, answers were counted as correct only if they were unprompted and occurred within 3 seconds from the time the questions were posed.

Experimental Procedure

Data on the performance of many other children who have participated in this program indicate that the children virtually never produce the target responses in settings where training has not occurred. Thus, correct sentences at school, correct paragraphs at school, correct sentences at home and correct paragraphs at home may be viewed as functionally unrelated response classes, or separate language tasks. The study employed a multiple-baseline design across these separate language tasks.

Baseline. During baseline on sentences and paragraphs at home, the teacher wrote a set of five questions and five answers for each child, and sent these home in the children's lunchboxes each school day. All questions pertained to children's experiences during that school day (see Table 2). When a child arrived home from school, a parent took the list of questions from the lunchbox, child and parent were seated opposite one another, and the parent proceeded to ask each question and to record the child's unprompted verbatim responses that occurred within 3 seconds.

If the child failed to respond within 3 seconds, the parent wrote "no response" on the data sheet and then verbally prompted the correct answer (e.g., "Say, 'Mr. Brown met me at the bus. Tommy was there, too!'") and praised the child for giving this prompted response. If the child made a response that was incorrect, the parent recorded it and then said "No" firmly, and prompted and praised the correct answer. These verbal prompts and behavior-specific praise for answers, were delivered throughout all baseline and treatment conditions.

During baseline on sentences and paragraphs produced at school, each child's parents wrote five questions and five answers about events that occurred at home during the preceding afternoon/evening, and sent these to school each day (see Table 3). Upon the child's arrival at school, the teacher took these questions from the lunchbox, teacher and child were seated opposite one another in a classroom, and the teacher asked the child each question and recorded the child's unprompted responses (or failures to respond) that occurred within 3 seconds. Like the parents, the teacher then used the verbal prompting procedure described above, as well as behavior-specific praise for both prompted and unprompted answers.

Baseline data collection on sentences and paragraphs at school began later than data collection on sentences and paragraphs at home. Although this results in an anomalous multiple-baseline design, it is in keeping with good parent-training strategy. Previous experience suggests that parents' ongoing involvement in this program is facilitated if they are not asked to begin writing questions and answers about children's home activities until shortly before the introduction of the behavior change procedure. Then, in the usual case, children's comparatively rapid acquisition of new verbal responses serves to maintain parents' active participation in the program.

Rehearsal. Sentences and paragraphs to be used at home were rehearsed at school, and sentences and paragraphs to be used at school were rehearsed at home. At school, rehearsal sessions were conducted in the afternoons, during children's regularly-scheduled 1:1 instructional sessions. Child 8's home rehearsals were conducted in the morning before he left for school, while Child 9's home rehearsals were conducted just before bedtime. Rehearsals usually lasted about 10 minutes.

During rehearsals, the teacher or parent asked the child the written questions about past events at school or home, and prompted correct answers as described earlier. An illustration of the rehearsal of a sentence to be used at home might be:

Teacher: Whose mommy was in school today?
Child: (No response)
Teacher: Say, "Bruce's mommy was in school."
Child: Brucie's mom was in school.
Teacher: Good, you said, "Brucie's mom was in school."

Paragraphs were rehearsed in a similar fashion. An example of the rehearsal of a paragraph to be used at school would be:

Parent: What game did you play with Jimmy when you got home from school?
Child: Tag. We play tag.
Parent: Say, "We played tag. It was fun."
Child: We played tag. It was fun.
Parent: Good for you; that's a right answer!

Each rehearsal session continued until the child could answer the questions correctly, within 3 seconds, and without prompts. Children earned tokens for their correct, unprompted responses, and exchanged these at the end of the rehearsal period for preferred foods, activities or toys.

Child 8 initially rehearsed two questions a day, and gradually moved to three, four and then five questions per day. Child 9, however, was always rehearsed on five questions per day.

Throughout all rehearsal conditions, questions continued to be asked and children's responses were recorded verbatim, as described earlier.

No Rehearsal. In this maintenance condition, rehearsals were discontinued for Child 9. As during baseline, parents and teacher continued to write sets of five questions and five answers about daily activities, and to send questions to the other setting. And using the same procedures as in baseline, parents and teacher continued to ask the child these questions and record her answers.

Reliability

A primary observer (the teacher) scored all written records of children's

responses to questions. A second observer, who was unfamiliar with the children and naive as to the purposes of the study, was also asked to score these transcripts. Interobserver agreement was obtained for at least 3 days (15 questions) in each baseline condition, for at least 5 days (25 questions) in each treatment condition, and for 5 days in Child 9's "no rehearsal" condition. Percent agreement was calculated using the formula: total number of agreements/total number of agreements + disagreements × 100.

Mean interobserver agreement for Child 8 was 100% during baseline and 97% during rehearsals (range = 92 to 100%). Mean interobserver agreement for Child 9 was 100% during baseline, 97% during rehearsals and 80% during the final no rehearsal condition (range = 80 to 100%).

Informal observation in the children's homes indicated that parents were correctly recording data and implementing the verbal rehearsal procedure. Further, both sets of parents frequently made notations on children's prompted responses, indicating that they were attending to the importance of the prompted versus unprompted dimension of the response definition. Two reliability estimates obtained in Child 9's home during the rehearsal condition for paragraphs yielded 80% interobserver agreement between parent trainer and father and 100% interobserver agreement between parent trainer and mother. Finally, it may be noted that parents were unaware of when verbal rehearsal procedures were begun at school.

RESULTS

Figures 8 and 9 display the children's answers to questions about temporally-remote (past) events that happened at school or home. During baseline, Child 8 answered only one question correctly (cf. paragraphs at school), and Child 9 answered none of the questions correctly.

When Child 8 (Figure 8) began school rehearsals to assist him in answering questions at home about events that happened during the school day, he initially practiced two questions. As he demonstrated his ability to answer questions correctly and in sentences on at least two consecutive days, additional questions were added, until he was able to answer five questions correctly and in sentences. Only five days after beginning rehearsals on paragraphs to be used at home, he answered five questions asked at home in paragraphs and subsequently, he averaged four paragraphic answers per day in reporting school events at home.

In the third month of the program, Child 8's parents began rehearsing him on home events to be reported at school. During 12 days of rehearsal, his use of paragraphs in reporting home activities and events at school increased from a baseline mean of zero to a mean of one, with a range of zero to two. Sentences at school are not reported for this child because the parents began home rehearsal of sentences before baseline data had been collected.

Child 9 (Figure 9) displayed even more substantial behavior change. After

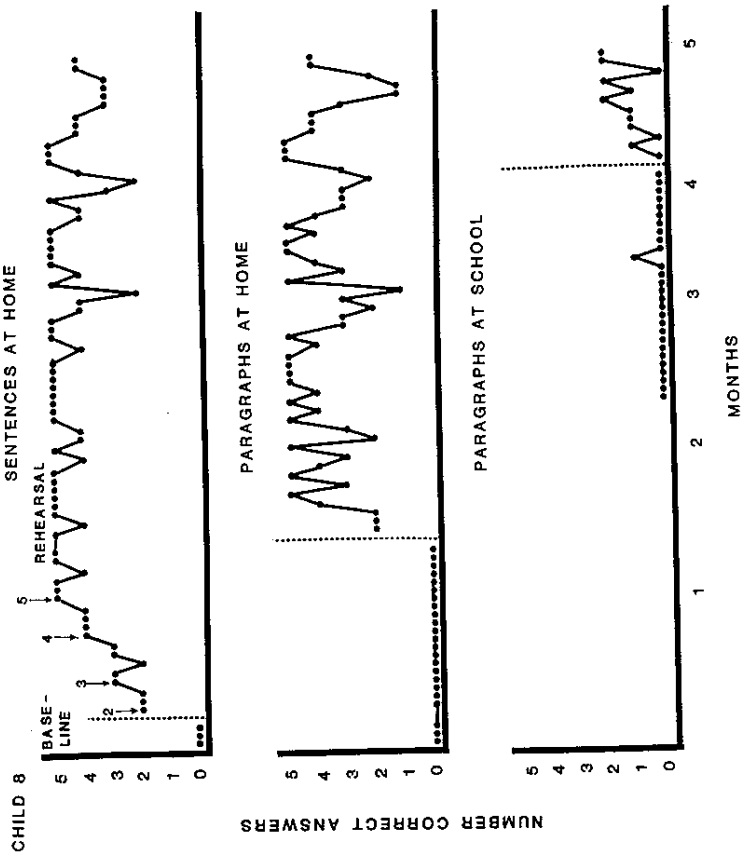

FIGURE 8. The number of correct sentence and paragraphic answers given by Child 8 to questions asked at home and at school, during baseline and rehearsal conditions. The number of questions asked after rehearsals were begun is indicated above the arrows.

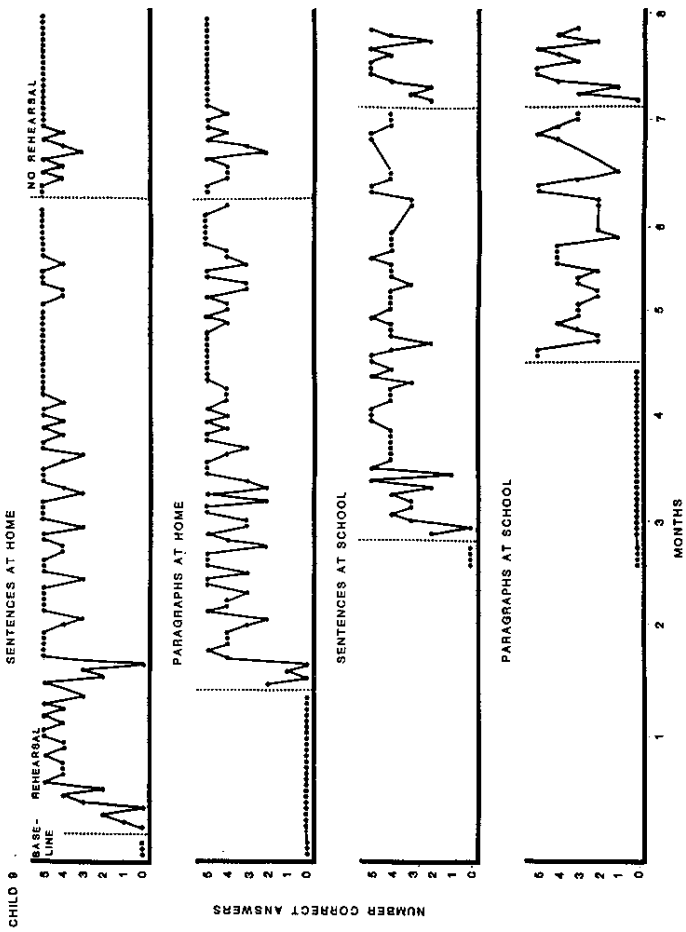

FIGURE 9. The number of correct sentence and paragraphic answers given by Child 9 to questions asked at home and at school during baseline, rehearsal and no rehearsal conditions.

school rehearsals of sentences were begun, her answers to questions asked at home increased from a zero baseline to a pattern of answering three to five questions in sentences, and by the end of the rehearsal condition, she was consistently answering four to five questions in sentences. School rehearsal of paragraphs to be used at home had a similarly rapid effect on her behavior—by the sixth school rehearsal, she was able to answer all five questions posed at home, using paragraphic speech.

Sentences rehearsed at home and used at school showed a similar increase following the initiation of home rehearsals; a mean of four questions asked at school was answered in a complete sentence. Finally, home rehearsal of paragraphs to be used at school resulted in a rapid change from a mean of zero during baseline to a mean of three questions answered correctly in paragraphs during this rehearsal condition.

When school rehearsals were discontinued early in the seventh month of intervention, Child 9 continued to answer home questions correctly and in paragraphs. During the last 16 days of the "no rehearsal" condition at school, all home questions were answered in paragraphs. Later deletion of home rehearsals resulted in a temporary decline in performance, but on the fifth day after home rehearsals were discontinued, the child achieved five correct paragraphic answers to questions written at home and asked at school.

DISCUSSION

This study demonstrated that a verbal rehearsal procedure was effective in increasing the accuracy and length of two autistic children's answers to questions about temporally-remote events. Rehearsal sessions often provided the teacher and parents with opportunities to correct grammatical, syntactical and semantic errors as well, since the intervention procedure facilitated this type of individualized instruction.

The teaching procedure reported above is readily adapted for use with children who exhibit more severe language deficits. For some children, verbal rehearsals may initially deal with correct answers in words or phrases, and later proceed to the rehearsal of sentences and paragraphs. In other cases (cf. Child 8) a gradually-increasing number of questions rehearsed has constituted a sufficient shaping procedure to help children achieve correct answers in sentences and later, in paragraphs. And with still other children, a "standard set" of questions has been initially introduced (e.g., "What did you eat for lunch?", "What song did you sing in music class?", "Who met you at the bus?") and rehearsals of different answers to the same questions have continued until children achieved success.

It is important to note that, although Children 8 and 9 were rehearsed on specific correct answers in one setting (home or school), the responses they made in the other setting were sometimes different from the rehearsed answers, al-

though accurate. The program appears to encourage generative speech, because memorization of a rehearsed answer is not the sole criterion for a correct, reinforced response, and because children are rehearsed on many different questions about remote events. Thus, during the "no rehearsal" condition for paragraphs at school, Child 9 averaged three sentences per question, although the rehearsal procedure was no longer being implemented.

Another key aspect of the procedure is initial school rehearsal of questions to be answered at home. Usually, this means that only two to five hours elapse before children respond to the rehearsed questions. Later home rehearsal of questions to be answered at school often involves a much greater time differential (sometimes as much as 12 hours), and many children appear to need practice in retaining information for shorter time periods before they attempt recall tasks about more temporally-remote events.

The intervention procedure reported here has been very popular with the families of autistic children. Twenty children and their parents have participated in this program over the past six years, and in the course of implementing the program, many parents have had their first real social conversations with their autistic child. In addition, parents who have expressed concern about their autistic child's failure to convey information about the events of the day have used this program as a vehicle for daily information gathering about the child's school experiences.

If, indeed, children's difficulties in recalling temporally-remote events are related to their language deficits, then language practice with regard to such events should be helpful in remediating these deficits. In the future, further research on the effects of immediate versus delayed rehearsals, and standard versus varied sets of questions should help to specify the parameters of the program that are salient for children who exhibit varying levels of language skills.

SUMMARY

Most autistic children require continued language intervention over a period of years. In the typical case, a child enters treatment with extremely minimal or nonexistent verbal skills, and training begins with nonverbal and later, verbal imitation tasks that contribute to phonological development and eventually, to functional words (usually, labels for people and objects). Subsequently, children learn common verbs, possessive pronouns, some "formula"-type sentences (e.g., "I want juice" or "I want top"). Later, they learn abstractions such as size/shape and color, and other concepts such as "yes" and "no," that enable them to engage in simple conversations (cf. Guess, Sailor, & Baer, 1976; Lovaas, 1977). This language instruction sequence, with some minor variations, is presently widely used in programs for autistic children throughout the country.

Historically, behavioral analyses of the language characteristics of autistic

and severely-retarded children began "at the beginning," with mute or severely-language-impaired children (cf. Risley, 1966; Baer, Peterson, & Sherman, 1968; Guess, Sailor, Rutherford, & Baer, 1968). Over the past 15 decades, a number of effective procedures have been identified to help children acquire more advanced language skills. In the meantime, however, two clinical trends may be noted. First, yesterday's autistic preschoolers are today's older adolescents or young adults, many of whom remain in treatment and need continued language programming to enhance the length and complexity of their verbal productions. Secondly, young autistic children entering treatment more recently often appear to benefit from improvements in instructional technology, so that they more quickly display readiness for programs that promote complex speech. Thus, there is a growing need for language intervention strategies that target those children who have attained more "advanced" language levels.

The three investigations reported above resulted in acquisition of additional language skills that contribute to the normalization of children's verbal productions, by facilitating more lengthy, complex and diverse expressive speech. Over the past three years, four children have participated in the multiple descriptor program and twelve children participated in the wh-concepts program. Over the past six years, 20 children and their families have participated in the remote events program. Thus, the results have been replicated many times.

While this research includes encouraging findings vis-à-vis response generalization across persons, settings and stimulus materials, it has been noted that it is easier to *establish* new language skills than to teach their *spontaneous use* in untrained settings (Guess, Keogh, & Sailor, 1978). Although the skill acquisition reported for Children 1 through 9 may eventually support the development of normative conversational speech, the investigation of procedures that facilitate complex, generative language continues to be an important area for future research.

Acknowledgement—The authors wish to express their appreciation to the many parents who have skillfully participated in providing language intervention services to their children.

REFERENCES

Baer, D. M. & Guess, D. Receptive training of adjectival inflections in mental retardates, *Journal of Applied Behavior Analysis*, 1971, **4**, 129–139.

Baer, D. M., Peterson, R. F. & Sherman, J. A. The development in imitation by reinforcing behavioral similarity to a model. *Journal of the Experimental Analysis of Behavior*, 1968, **10**, 405–416.

Boucher, J. & Warrington, E. K. Memory deficits in early infantile autism: Some similarities to the amnesic syndrome. *The British Journal of Psychology*, 1976, **67**, 73–87.

Brown, R. The development of Wh questions in child speech. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 279–290.

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

Fay, W. H., & Schuler, A. L. *Emerging language in autistic children.* Baltimore: University Park Press, 1980.

Frank, S. M., Allen, D. A., Stein, L., & Myers, B. Linguistic performance in vulnerable and autistic children and their mothers. *American Journal of Psychiatry*, 1976, **133**, 909–915.

Gallagher, T. M. & Darnton, B. A. Conversational aspects of the speech of language-disordered children: revision behaviors. *Journal of Speech and Hearing Research*, 1978, **1**, 118–135.

Garcia, E., Guess, D., & Byrnes, J. Development of syntax in a retarded girl using procedures of imitation, reinforcement, and modeling. *Journal of Applied Behavior Analysis*, 1973, **6**, 299–310.

Guess, D. A functional analysis of receptive language and productive speech: acquisition of the plural morpheme. *Journal of Applied Behavior Analysis*, 1969, **2**, 55–64.

Guess, D., Keogh, W., & Sailor, W. Generalization of speech and language behavior: measurement and training tactics. In R. L. Schiefelbusch (Ed.), *Bases of language intervention.* Baltimore: University Park Press, 1978.

Guess, D., Sailor, W., & Baer, D. M. *Functional speech and language training for the severely handicapped.* Lawrence, Kansas: H & H Enterprises, Inc., 1978.

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

Handleman, J. S. Generalization by autistic-type children of verbal responses across settings. *Journal of Applied Behavior Analysis*, 1979, **12**, 273–282.

Harris, S. L. *Teaching speech to a nonverbal child.* Lawrence, Kansas: H & H Enterprises, Inc., 1976.

Hart, B. M. & Risley, T. R. Establishing use of descriptive adjectives in the spontaneous speech of disadvantaged preschool children. *Journal of Applied Behavior Analysis*, 1968, **1**, 109–120.

Hart, B. & Risley, T. R. Using preschool materials to modify the language of disadvantaged children. *Journal of Applied Behavior Analysis*, 1974, **7**, 243–256.

Hermelin, B. Rules and language. In M. Rutter (Ed.), *Infantile autism: Concepts, characteristics, and treatment.* London: Churchill Livingstone, 1971.

Hermelin, B. Coding and the sense modalities. In L. Wing (Ed.), *Early childhood autism.* London: Pergamon Press, 1976.

Heward, W. L., & Eachus, H. T. Acquisition of adjectives and adverbs in sentences written by hearing impaired and aphasic children. *Journal of Applied Behavior Analysis*, 1979, **12**, 391–400.

Lahey, B. B. Modification of the frequency of descriptive adjectives in the speech of Head Start children through modeling without reinforcement. *Journal of Applied Behavior Analysis*, 1971, **4**, 19–22.

Lovaas, O. I. *The autistic child: Language development through behavior modification.* New York: Irving Publishers, Inc., 1977.

Lutzker, J. R., & Sherman, J. A. Producing generative sentence usage by imitation and reinforcement procedures. *Journal of Applied Behavior Analysis*, 1974, **7**, 447–460.

Maloney, K. B., & Hopkins, B. L. The modification of sentence structure and its relationship to subjective judgements of creativity in writing. *Journal of Applied Behavior Analysis*, 1973, **6**, 425–433.

Martin, J. A. Generalizing the use of descriptive adjectives through modelling. *Journal of Applied Behavior Analysis*, 1975, **8**, 203–209.

Quirk, R., & Greenbaum, S. *A concise grammar of contemporary English.* New York: Harcourt Brace Jovanovich, Inc., 1973.

Risley, R. R. *The establishment of verbal behavior in deviant children.* Ph.D. dissertation, University of Washington (Seattle). 1966.

Risley, T. R., & Wolf, M. M. Establishing functional speech in echolalic children. *Behavior Research and Therapy*, 1967, **5**, 73–88.

Sailor, W. Reinforcement and generalization of productive plural allomorphs in two retarded children. *Journal of Applied Behavior Analysis*, 1971, **4**, 305–310.

Sailor, W., & Taman, T. Stimulus factors in the training of prepositional usage in three autistic children. *Journal of Applied Behavior Analysis*, 1972, **5**, 183–190.

Schreibman, L., & Carr, E. G. Elimination of echolalic responding to questions through the training of a generalized verbal response. *Journal of Applied Behavior Analysis*, 1978, **11**, 453–463.

Siegel, G. M., & Spradlin, J. E. Programming for language and communication therapy. In R. L. Schiefelbusch (Ed.), *Language intervention strategies*. Baltimore: Unviersity Park Press, 1978.

Sosne, J. B., Handleman, J. S., & Harris, S. L. Teaching spontaneous-functional speech to autistic-type children. *Mental Retardation*, 1979, **17**, 241–246.

Stevens-Long, J., & Rasmussen, M. The acquisition of simple and compound sentence structure in an autistic child. *Journal of Applied Behavior Analysis*, 1974, **7**, 473–479.

Stevens-Long, J., Schwarz, S. & Bliss, C. The acquisition and generalization of compound sentence structure in an autistic child. *Behavior Therapy*, 1976, **7**, 397.

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

Twardosz, S., & Baer, D. M. Training two severely retarded adolescents to ask questions. *Journal of Applied Behavior Analysis*, 1973, **6**, 665–661.

Warren, S. F., & Rogers-Warren, A. Current Perspectives in Language Remediation. *Education and Treatment of Children*, 1980, **3**, 133–152.

Wilcox, J. M., & Leonard, L. B. Experimental acquisition of wh-questions in language disordered children. *Journal of Speech and Hearing Research*, 1978, **4**, 220–239.