JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1986, 19, 289–297 NUMBER 3 (FALL 1986)

SOME DIRECT AND GENERALIZED EFFECTS OF REPLACING AN AUTISTIC MAN'S ECHOLALIA WITH CORRECT RESPONSES TO QUESTIONS

MARTIN J. MCMORROW AND R. M. FOXX

ANNA MENTAL HEALTH AND DEVELOPMENTAL CENTER

We extended the use of operant procedures to decrease immediate echolalia and increase the appropriate responding to questions of a 21-year-old autistic man. Three experiments were conducted in which the overall plan was to (a) encourage the subject to remain quiet before, during, and after the presentation of questions and (b) teach him to use environmental cues (i.e., word cards or a model's responses) to increase the likelihood of responding correctly. Multiple baseline designs demonstrated that echolalia was rapidly replaced with correct stimulus-specific responses. In addition, there were a variety of generalized improvements in the subject's verbal responses to questions. The procedures and results are contrasted to previous research in an attempt to explain the encouraging findings.

DESCRIPTORS: autism, echolalia, language training, operant procedures, generalization, prompts

Immediate echolalia (i.e., repetition of one or more words in a temporally-related sample verbalization) is a common language characteristic of autistic individuals that can interfere with communication, learning, and social development. Although theories exist regarding the development and function(s) of echolalia (e.g., Prizant & Duchan, 1981; Rutter, 1978), few programs to teach echolalics more adaptive communication have been empirically evaluated.

Thus far, the most effective programs to teach correct stimulus-specific verbal responses to echolalics have used operant procedures such as imitation training, stimulus fading, and differential reinforcement (e.g., Carr, 1985; Cushing, Adams, & Rincover, 1983; Harris, 1975). For example, Carr, Schreibman, and Lovaas (1975) and Risley and Wolf (1967) taught correct responses to questions by having the trainer (a) present the verbal stimulus (e.g., "What is a rose?"); (b) follow it with a verbal prompt that was likely to be echoed (e.g., "flower"); (c) differentially reinforce the imitative response; and (d) fade the verbal prompt once it was reliably echoed. Both studies produced increases in correct responding and decreases in immediate echolalia to trained verbal stimuli.

However, the clinical practicality of this method has been questioned because there has been little, if any, change in correct responding or echolalia in response to novel stimuli (i.e., untrained questions). Findings such as these led Carr et al. (1975) and Schreibman and Carr (1978) to conclude that reducing echolalia by teaching appropriate responses is not practical clinically because it would be impossible to teach the echolalic individual a response to every verbal stimulus that might be encountered.

In view of this limitation, recent studies have taken different approaches with echolalia. For example, Schreibman and Carr (1978) taught echolalics to respond "I don't know," to a small set of previously echoed questions and produced generalization to a broad set of previously echoed questions without affecting their nonecholalic responses. In contrast, Charlop (1983) used echolalia to facilitate the acquisition and generalization of nonverbal appropriate behavior (i.e., receptive labeling). Although these approaches have yielded encouraging results, the problem inherent in the pioneering work remains; namely, how to replace

We wish to thank Jayne Ness, Sarah Fenlon, Carolyn Bird, and Leon Rendleman for their assistance.

Requests for reprints or complete lists of the trained questions and correct responses may be sent to either author at the Anna Mental Health and Developmental Center, 1000 N. Main St., Anna, Illinois 62906.

echolalia with stimulus-specific responses in a way that is clinically practical and produces generalized improvements in verbal communicative behavior.

Our research attempted to address this problem. Our overall goal was to teach an echolalic subject to remain quiet before, during, and after the presentation of verbal stimuli and then make stimulus-appropriate verbalizations by using environmental cues. Three experiments were conducted. The first used mild verbal reprimands, external cues, manual prompts, response-specific feedback, reinforcement, and stimulus fading in a highly structured training situation. The second and third experiments used reinforced modeling, feedback, and reinforcement in an attempt to accelerate learning, teach the subject to attend to others' responses, and mediate longer latencies between the presentation of the stimuli and his opportunity to respond.

METHOD

Subject and Settings

Tom was a 21-year-old man who had been at our facility for 6 months. He was diagnosed as mentally retarded (IQ of 40, Peabody Picture Vocabulary Test), although he had previously been diagnosed as autistic and exhibited all of the characteristics of autism described by Schreibman (1975). Tom's verbal behavior was limited in a number of ways. First, his instructors and mother could not recall him ever verbally initiating an interaction. In fact, his only nonverbal "initiations" were to approach staff during token exchange and shower times (i.e., for points or a towel). Second, although Tom rarely failed to respond verbally to verbal stimuli (e.g., statements or questions), he virtually always repeated one or more of the words and his responses almost never included any other words. Finally, he rarely used yes or no responses to choice questions (e.g., "do you want ____?") and, if so, at only the level of chance. Despite these limitations, Tom had a few skills that we attempted to use. Most importantly, he could verbally label a variety of pictures, objects, and printed words when verbally prompted. Interestingly, the verbal stimuli commonly used in labeling exercises (e.g., "What is this?") were the only ones he did not echo. In addition, Tom could follow many simple instructions that did not require expressive language.

All training was conducted in a 7 m × 9.5 m room that contained three tables and several chairs. The generalization tests in Experiment I were conducted in a small lounge in a different part of the building.

Target Behaviors, Recording, and Reliability

In each experiment, the first word or sequence of words that followed a question was scored in one of three mutually exclusive categories: echolalia, incorrect, or correct. Echolalia was scored when Tom repeated one or more of the words in the question regardless of whether other verbalizations occurred. An incorrect response was scored when his response contained an irrelevant word(s), regardless of whether a correct response was included. A correct response was scored when his verbalization either matched the trained response or provided a different appropriate answer. Using these definitions, a response that contained any combination of echolalic, correct, or incorrect verbiage was scored as an echo and a response that included any combination of correct and incorrect verbalizations was scored incorrect. We used these strict definitions for three reasons. First, since Tom's verbalizations rarely included more than one word, we suspected they could be easily categorized. Second, the pervasiveness and one word nature of Tom's echolalia suggested that he was not using it in any mediational or functional way (e.g., Charlop, 1983). Third, we planned to use feedback procedures that were specific to each type of response in order to increase the likelihood of generalized improvements in Tom's language behavior.

Responses were scored by the trainer immediately after each question. All sessions were audiotaped. An independent rater transcribed and scored a random 20% of the audiotapes from each condition in each experiment and all of the generaliza-

tion assessments in Experiment I. Excluding the few taped verbalizations that were inaudible, interrater agreement on all target behaviors was always 100%.

EXPERIMENT I

Thirty questions from three content areas were targeted for training (i.e., 10 questions per area). The *identification* area included questions such as "What is your name?" and "Where do you live?" *Interaction* questions included "How are you today?" and "What kind of music do you like?" whereas *facts/figures* questions included "What state do you live in?" and "What baseball team plays in St. Louis?" There were from one to four words in the trained responses. Tom's responses to these questions had not been assessed prior to the study.

Experimental Design

Training was implemented in a multiple baseline design across content areas. From three to four trials (i.e., the presentation of all 10 questions in a content area) were conducted daily on each set during sessions that lasted approximately 30–40 min. The generalization tests in Experiment I were conducted by a uniformed institutional security officer, a peer, and an unknown staff member prior to baseline and following training. Tom was given no feedback regarding his responses during these tests and no training personnel were present.

Procedures

Response identification training. Prior to Experiment I, an assistant taught Tom to verbally label words that would be trained as correct responses. Each response (i.e., word or set of words) was hand-lettered on an index card. Labeling training consisted of randomly presenting the cards verbally (i.e., "what does this say?") and gesturally (i.e., pointing to or tapping the card), prompting Tom to identify the words on the card, providing verbal feedback, saying the correct response when Tom failed to do so and then prompting him (as above) to repeat it, and giving praise and sips of soda for each correct response. Training continued until Tom identified all of the responses in each area during three consecutive trials when the trainer simply pointed to the card. Each set was trained separately. Eleven to 15 trials were required to achieve criterion and the total training time ranged from 15 to 25 min per set.

Baseline. The baseline was conducted after Tom had completed response identification training. Three tables were used so that the 10 response cards from each content area could be placed on a different one. Soda and a transparent cup also were placed on the table. After Tom was seated, the trainer (first author) sat across from him and said, "I am going to ask you some questions and I want you to answer them the best you can." He then asked the 10 questions in a random order and provided specific feedback following each response (i.e., "that's right" for a correct response, "that's not right" for an incorrect response, and a gentle "no" for an echo). Tom received a sip of soda for each correct response. After a trial on one set was completed, Tom and the trainer simply moved to the next table and began another set. Response-specific feedback and reinforcement were provided during baseline because (a) they would be used during training and it was necessary to demonstrate that they would not have been effective alone, (b) we felt they represented a strategy that might be used by others (e.g., teachers or parents), and (c) we hoped that these procedures might influence Tom's responding following training.

Training phase I (cards, pause, point). This training taught Tom to remain silent until the trainer prompted him to label one of the response cards on the table by having the trainer (a) hold up his right index finger at eye level midway between Tom and himself whenever silence was desired (i.e., during the instructions, questions, and for approximately 1 s following the question) and say "no" or "shh" whenever Tom verbalized (i.e., the pause prompt); (b) move his finger so that it touched the correct response card approximately 2 s after the question was completed (recall that at the end of response identification training Tom was labeling to the point prompt only); (c) use the

response identification training prompts (i.e., tapping the card, "what does this say?"), if necessary to ensure that the labeling response occurred; (d) cover the response card with his right hand and use a bridging stimulus (e.g., a head nod or smile) immediately following Tom's identification of the word(s) on the card; (e) raise his left index finger to eye level (i.e., pause prompt again), restate the question, and move this finger so that it touched the back of his right hand when a correct response was desired (i.e., he point prompted again and used the response identification prompts if necessary even though the response card was covered); and (f) provide the same verbal feedback and consequences that were used in baseline for the first verbalization that occurred. The trainer continued to randomly present the questions as he had in baseline. Tom's response when the card was covered was the one that was scored.

Probe. A return-to-baseline probe was conducted on each set to determine if Tom would use the response cards when no prompts were used.

Training phase II (pause only). The cards were removed and no point prompts were used. The trainer simply pause prompted as he asked questions and then lowered his hand so that it rested in front of him when it was time for Tom to respond. Feedback and reinforcement were provided as before.

Baseline II. This condition was the same as the initial baseline except that the cards were absent. Thus, the trainer simply asked each question and provided feedback and consequences as before.

Programming generalization and maintenance. This condition consisted of several phases. First, sessions were conducted as during baseline II by a new trainer (T2). Second, the original trainer returned and faded the feedback and consequences by progressively reducing the number of words used to praise and providing sips of soda for every other correct response before gradually eliminating all feedback and consequences. Third, three different individuals questioned Tom in the training room but provided no feedback or consequences. Finally, the postgeneralization assessments were conducted by the individuals who had conducted the preassessments.

Follow-up. One month after the postgeneralization assessments, the original trainer simply asked Tom all of the questions with no response cards, prompts, feedback, or consequences being used. It is important to note that Tom had received training on other question and response sets (Experiment II) during this 1-month period.

Results

The mutually exclusive scoring procedures had little, if any, influence on the data. For example, responses that contained echolalic and correct verbalizations (i.e., Tom echoing and then producing a correct response or responding correctly and then echoing) occurred after only 19 of the 2,810 questions.

Figure 1 shows that Tom echoed in response to 70%–100% of the questions during baseline even though the response cards were present and response-specific feedback was provided. He responded correctly to only 3 of the 30 questions across the three sets, and repeated exposures with feedback did not increase his correct responding. The addition of the pause and point prompts replaced echolalia with 100% correct responding during the third training trial on the identification set and during the second trial on the interaction and facts/figures sets. The return-to-baseline probe across all three sets (trial 25) suggested that the pause and point prompts, not the response cards, were responsible for the decreases in echolalia and increases in correct responding at that point in training. As a result, additional training trials were conducted during which near-errorless responding occurred and training phase II was planned. During training phase II (i.e., pause only), Tom's correct responding initially decreased on each set but recovered to a mean of between 70% and 100% correct on each set during the final three trials. Of particular interest was that echoing remained low in the early trials and disappeared in the last 14 trials of the condition even though Tom was not always responding correctly. Tom was now using previously reinforced responses incorrectly (i.e., virtually all incorrect responses were misuses of responses trained for other questions). Thereafter, correct responding was high, echolalia was absent,

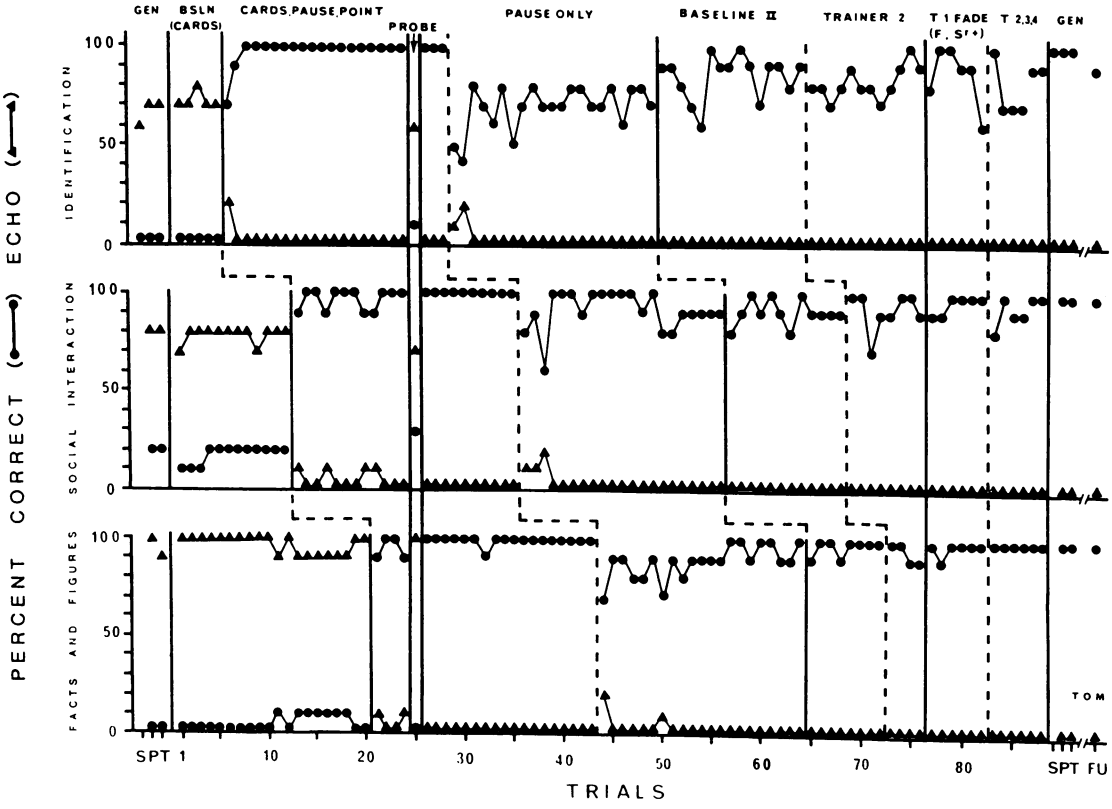

Figure 1. Experiment I. The percentage of echolalic and correct responses to questions in three content areas. The S, P, and T labels during the "Gen" (i.e., generalization) phases denote the security officer, peer, and staff member, respectively. In the T1 FADE condition, the "F" and "S^{r+}" denote feedback and positive consequences. T2, 3, 4 refer to trials in which persons other than the primary trainer presented the questions. The follow-up (F.U.) was conducted 1 month after the postgeneralization assessments.

and errors were the result of misusing correct responses to other questions.

During the pregeneralization trials, Tom echoed 60%–100% of the questions asked by the security officer, peer, and staff member and never answered more than 2 of the 30 questions correctly; during the postassessments he never echoed and answered all of the questions correctly. In addition, he responded correctly to 90%–100% of the questions at the 1-month follow-up. These results were very encouraging given that he had received no additional exposure to these questions.

EXPERIMENT II

Our next goal was to teach Tom to attend to the correct responses of another person and rely on them to answer questions correctly. Experiment II began the day after the postgeneralization tests in Experiment I.

The trainer, setting, and data collection procedures were the same as in Experiment I and a multiple baseline design was used. Two new sets of 10 questions were developed (e.g., "What is your favorite game?", "What month is Christmas in?") but only three conditions were run. Response identification training was not used; thus, Tom was not exposed to the correct responses until training started.

Baseline. This condition was identical to baseline II in Experiment I. The trainer simply presented the questions and provided feedback and consequences dependent on Tom's verbal responses.

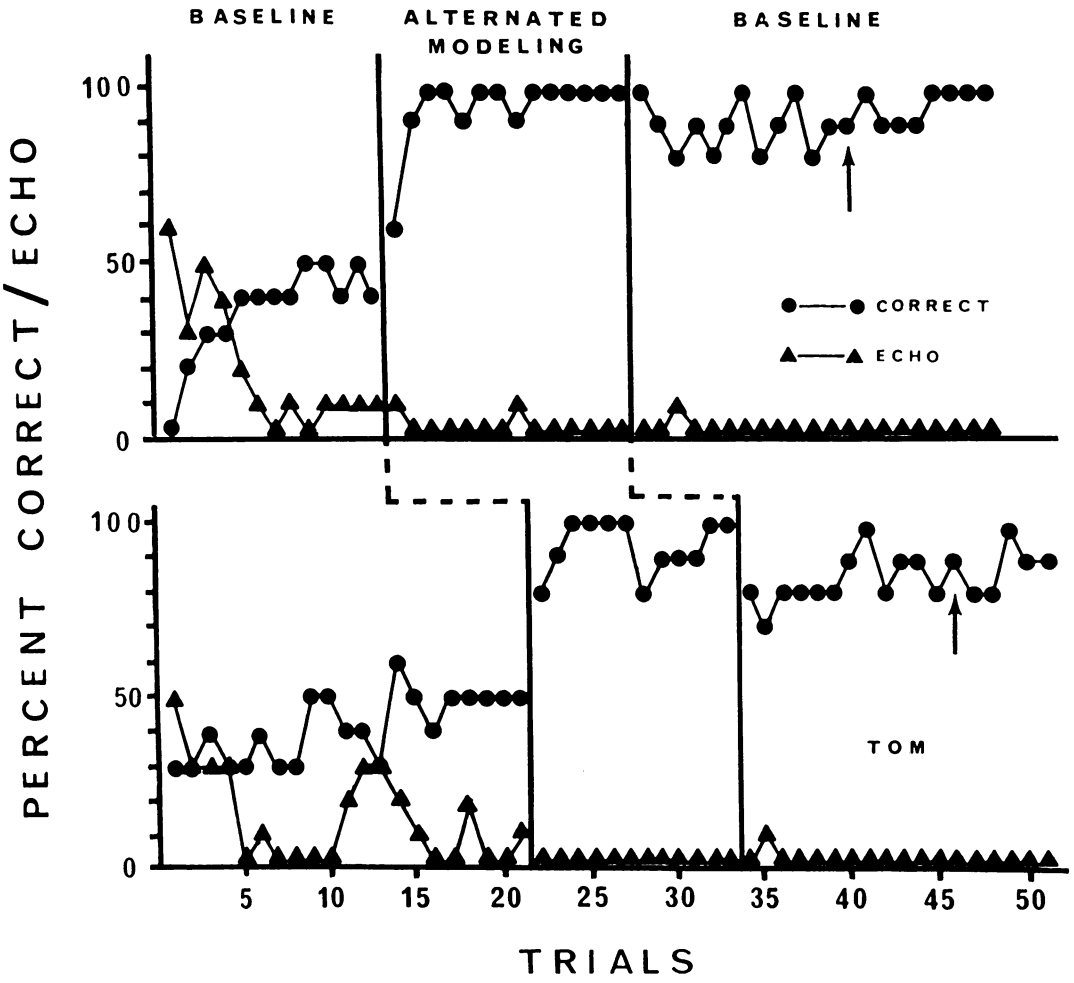

Figure 2. Experiment II. The percentage of echolalic and correct responses to questions in two response sets. The arrows show where feedback and consequence fading procedures were initiated.

Alternated modeling. Tom and a model sat across the table from the trainer. The trainer (a) looked at the model, asked a question, and provided the feedback and positive consequence for the correct response (i.e., "good answer," and a sip of soda), and then (b) looked at Tom, asked the same question, and provided the same feedback and consequence. This alternated sequence (i.e., ask the model, then ask Tom) was continued in each trial until all 10 questions had been presented. As before, the questions were randomly presented each trial.

Baseline. The model was removed and, after 12 trials on each response set, feedback and consequences were faded and then eliminated in two or three trials as in Experiment I.

Results

In marked contrast to the initial baselines in Experiment I, Figure 2 shows that Tom's echolalia and correct responding were influenced by the feedback and consequences used during baseline. Consider that on set 1, correct responding increased from zero in the first trial to between 40% and 50% correct in the last nine trials, while echoing decreased from 60% in trial one to between zero and 10% during the last eight trials. Similar results were obtained on set 2. These results were inter-

esting because no planned environmental cues were present that could have influenced correct responding and Tom was responding incorrectly more often than echoing. During the alternated modeling condition, Tom reached 100% correct on both sets within three trials and his echoing was virtually eliminated. Thereafter, his correct responding was never lower than 80%. When the model was withdrawn during the return to baseline, correct responding showed little change.

EXPERIMENT III

Since Experiment II demonstrated that Tom could use another's verbal behavior appropriately (instead of echoing it), our next goal was to expand this skill by building in a delay and competing stimuli between the modeled responses and Tom's opportunity to respond. Experiment III began the day after Experiment II ended and was identical to Experiment II except that two new sets of questions were used and, during a *continuous modeling* condition, the trainer asked the model all 10 questions before presenting them to Tom.

Results

Figure 3 shows that during the feedback and consequences baseline, Tom responded in much the same way as he had during the baseline of Experiment II. Towards the end of baseline Tom was responding incorrectly, rather than echoing, to almost half of the questions across the two sets. During the continuous modeling condition, Tom's correct responding reached 100% in the tenth training trial on both sets. Although Tom had never previously taken more than three trials to reach 100% correct, it seemed reasonable that it would have taken him longer, given the delay and competing stimuli between the modeled correct responses and his opportunity to respond. Echoing occurred only once during the last 11 training trials on both sets. In the return to baseline, Tom's correct responding remained at 100% in set 1 and varied from 80% to 100% in set 2. Echoing occurred only once during the entire condition and a high level of correct responding was maintained when the feedback and consequences were removed.

DISCUSSION

Experiment I supported previous research (e.g., Carr et al., 1975; Risley & Wolf, 1967; Schreibman & Carr, 1978) demonstrating that operant procedures can be used to increase appropriate responses and reduce echolalia following trained verbal stimuli. It also provided some support for the contention that replacing echolalia with stimulus-specific responses may be of limited clinical practicality (e.g., Carr et al., 1975). For example, the cards-pause-point training on set 1 had little, if any, positive effect on Tom's responding to questions in the baselines of sets 2 and 3 even though feedback and positive consequences for correct responses were used. In other words, no immediate spread of effect resulted from the training on set 1.

Other data from the three experiments, however, suggest that the cards-pause-point training may have produced several desirable effects on Tom's responses to untrained stimuli. Most importantly, during the initial baselines in Experiments II and III, Tom's correct responding increased and echoing decreased, even though feedback and consequences had previously been ineffective during the first baselines in Experiment I. Second, correct responding increased during the initial baselines of Experiments II and III, even though no planned environmental cues (e.g., response cards) were present. Third, correct responding either improved (following initial decreases) or was maintained across posttraining conditions in which prompts, cues, feedback, and/or consequences were being eliminated. Fourth, echolalia rarely occurred after Tom had been exposed to correct responses, even though he did not always respond correctly. Fifth, the training became easier, less directive, and less time-consuming across the three experiments.

There are several possible explanations why our findings differ from previous research. First, they could be specific to this subject. Clearly, replications will be needed to determine the across-subject

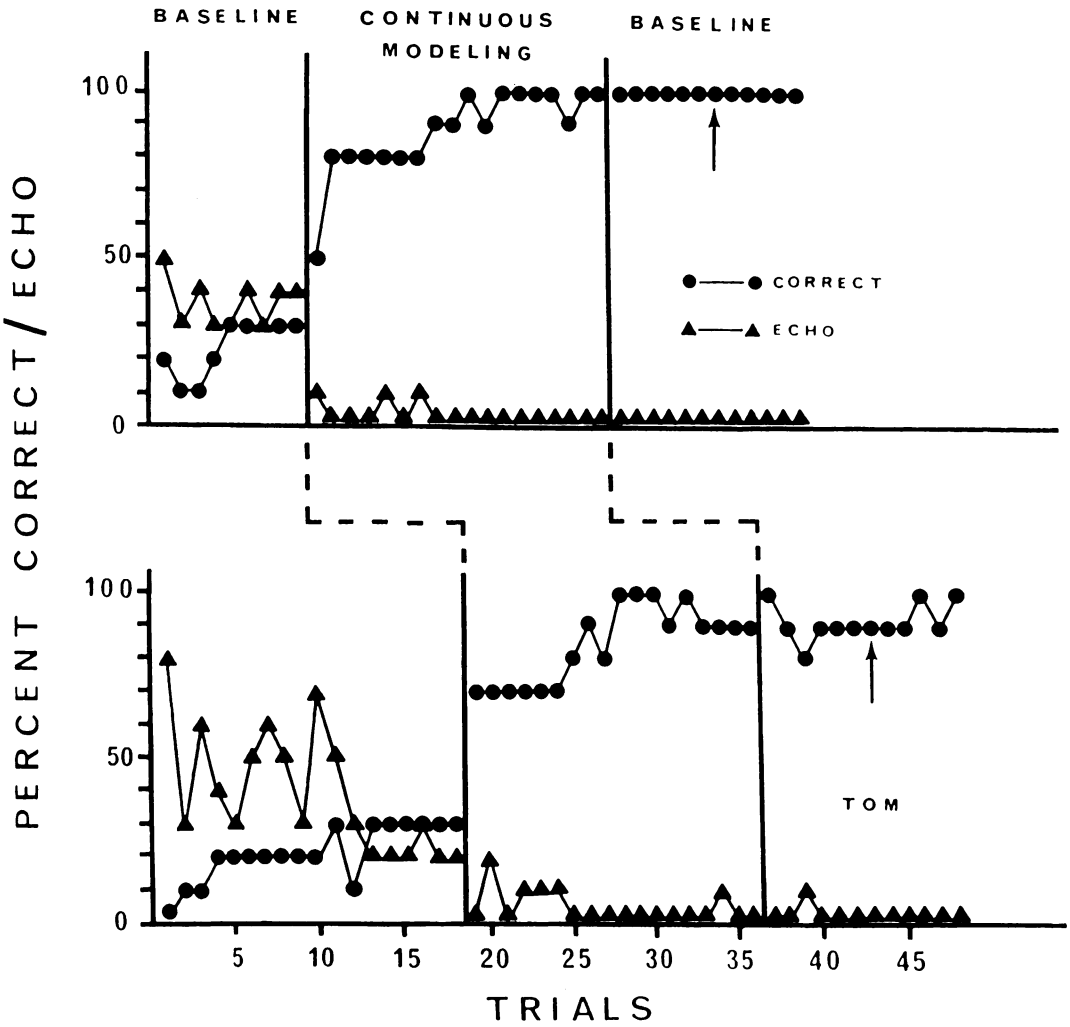

Figure 3. Experiment III. The percentage of echolalic and correct responses to questions in two response sets. The arrows show where feedback and consequence fading procedures were initiated.

generality of these effects. Second, because previous research has generally trained fewer responses and assessed generalization with fewer untrained stimuli for shorter periods of time, it could be that generalized effects may not occur until a particular number of exemplars have been trained or a subject-specific amount of training has been conducted. Indeed, had we not run a second experiment that created an opportunity to observe Tom in a new baseline condition with untrained stimuli, our conclusions may have more closely resembled those of previous research. Third, procedural factors may have accounted for the differences. For example, it could be that earlier studies (e.g., Carr et al., 1975; Risley & Wolf, 1967) did not obtain improvements on untrained stimuli because echoing on trained questions was reinforced prior to eliminating the verbal prompts (i.e., the desired responses that the subject was likely to echo). In contrast, we never reinforced echolalia and, at times, actively discouraged it with prompts and gentle reprimands. Fourth, our effects may have been related to the use of manual rather than verbal prompts that encouraged Tom to attend to environmental cues. For example, the pause and point prompts helped ensure that Tom remained quiet during the

instructions, attended to the trainer during the questions, and delayed his verbal response until the trainer had time to indicate the appropriate environmental cue (i.e., point to the correct response card). In addition, the trainer's control over Tom's verbalizations was conveniently enhanced because the end of the pause prompt (i.e., one finger up) also constituted the beginning of the point prompt.

To us, the cards-pause-point procedures not only taught correct responding and reduced echolalia, but produced other behavior changes that enhanced generalization. For example, Tom began to delay his responding over time in conditions where no prompts were used. This suggests that learning to pause reduced his echolalia. In addition, he apparently learned to use environmental cues to generate correct responses since they sometimes increased in the absence of prompts. For example, during training in Experiments II and III Tom often hesitated after a question, looked at the model, and then gave the correct response. This suggested that he learned a communicative strategy wherein echolalia never resulted in reinforcement while other responses, particularly correct ones, were encouraged.

Nevertheless, our enthusiasm is tempered by the knowledge that more systematic research is needed to (a) determine the external validity of these results; (b) experimentally isolate if, when, why, and to what extent generalized effects may occur; and (c) assess the relative, sequential, or combined effects of the various procedure components. Perhaps most importantly, research is needed to determine whether the procedures could be used with other types of environmental cues (e.g., objects) since this would greatly enhance their clinical applicability.

REFERENCES

Carr, E. G. (1985). Language acquisition in developmentally disabled children. *Annals of Child Development,* **2,** 49–76.

Carr, E. G., Schreibman, L., & Lovaas, O. I. (1975). Control of echolalic speech in psychotic children. *Journal of Abnormal Child Psychology,* **3,** 331–351.

Charlop, M. H. (1983). The effects of echolalia on acquisition and generalization of receptive labeling in autistic children. *Journal of Applied Behavior Analysis,* **16,** 111–126.

Cushing, P., Adams, A., & Rincover, A. (1983). Research on the education of autistic children. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (pp. 1–48). New York: Academic Press.

Harris, S. L. (1975). Teaching language to nonverbal children—with emphasis on problems of generalization. *The Psychological Record,* **82,** 565–580.

Prizant, B. M., & Duchan, J. F. (1981). The functions of immediate echolalia in autistic children. *Journal of Speech and Hearing Disorders,* **46,** 241–249.

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

Rutter, M. (1978). Diagnosis and definition. In M. Rutter & E. Schopler (Eds.), *Autism: A reappraisal of concepts and treatment* (pp. 1–26). New York: Plenum.

Schreibman, L. (1975). Attentional deficits in autistic children: Implications for teaching. *Claremont Reading Conference, 39th Yearbook* (pp. 159–166). Claremont, CA.

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

Received October 21, 1985
Final acceptance April 16, 1986