JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1985, 18, 127-139 NUMBER 2 (SUMMER 1985)

TEACHING BETWEEN-CLASS GENERALIZATION OF TOY PLAY BEHAVIOR TO HANDICAPPED CHILDREN

THOMAS G. HARING

UNIVERSITY OF CALIFORNIA, SANTA BARBARA

In this study, young children with severe and moderate handicaps were taught to generalize play responses. A multiple baseline across responses design, replicated with four children, was used to assess the effects of generalization training within four sets of toys on generalization to untrained toys from four other sets. The responses taught were unique for each set of toys. Across the four participants, training to generalize within-toy sets resulted in complete between-class generalization in 11 sets, partial generalization in 3 sets, and no generalization in 2 sets. No generalization occurred to another class of toys that differed from the previous sets in that they produced a reaction to the play movement (e.g., pianos). Implications for conducting research using strategies based on class interrelationships in training contexts are discussed.

DESCRIPTORS: generalization, response class, severely handicapped, stimulus generalization, leisure skill activities

Although developmental theorists have described responses as occurring in organized systems (e.g., Piaget, 1980), and have indicated that the organization of responses may influence generalization (Husaim & Cohen, 1981), behavior analysts have only recently studied some of the possible effects of response interrelationships. Voeltz and Evans (1982) reviewed the existing literature concerning response interrelationships. In those studies reviewed, response interrelationships were usually defined as an alteration in the frequency of a response when the frequency of another response changed as a function of changes in environments or the addition of a treatment variable.

The construct of the response class (Skinner, 1935, 1953) has been invoked to theoretically account for observed interrelationships between responses (e.g., Sherman, 1964). Inherent in the definition of a response class is that responses of different form may occur under the same or similar stimulus conditions if the responses are effective in producing similar effects. Therefore, an alteration designed to affect a single response may also affect functionally related responses.

Two strands of research have contributed demonstrations of response-response relationships. A variety of statistical models have been used to identify clusters of responses including factor analysis (Kara & Wahler, 1977), cluster analysis (Lichstein & Wahler, 1976), and lag sequential analysis (Strain & Ezzell, 1978). Another research strategy has established an intervention oriented approach (e.g., Wahler, Sperling, Thomas, Teeter, & Luper, 1970). Within language training research, several studies (Guess & Baer, 1973; Lee, 1981; Whitehurst, 1977) have shown interrelationships between receptive and productive language acquisition. Inverse relationships have been found between behavior problems and more situationally appro-

Preparation of this article was supported in part by Grant #G008104154, Special Education Program, U.S. Department of Education. I thank John Healy and Blair Roger for their administrative support, Melanie Lee for conducting the study on a daily basis, and Cathy Breen for help in designing the procedures. Anonymous reviewer "F" made substantial and important comments on an earlier draft of this article.

I gratefully acknowledge the assistance I received from the members of my doctoral committee from San Francisco State University and the University of California, Berkeley: Eli Bower, Eileen Gambrill, and Wayne Sailor. I am indebted to Robert Gaylord-Ross who served as my dissertation advisor and mentor throughout the work on which this article is based.

Requests for reprints can be sent to Thomas Haring, Graduate School of Education, University of California, Santa Barbara, California 93106.

priate behaviors (e.g., Haring, Breen, Pitts-Conway, & Gaylord-Ross, 1984; Koegel & Covert, 1972; Russo, Cataldo, & Cushing, 1981). Although response interrelationships have frequently been documented when multivariate measurement strategies have been used, interrelationships are not an inevitable product of behavioral interventions (Neef, Shafer, Egel, Cataldo, & Parrish, 1983).

Given that many studies have found response interrelationships either as directly programmed effects or as unintended effects, the implication can be made that a technology to generate response interrelationships is possible. Although the effects of response interrelationships can be observed, there are few data concerning how response interrelationships are initially formed. It would be useful to know if procedures designed to facilitate the acquisition of new response-response relationships could be developed. Research that validated procedures that promote response class relationships would have considerable significance to applied research in that such methods offer the potential to increase the economy of behavioral interventions.

There has been little human research concerning the effects of response interrelationships on stimulus generalization, although Casalta (1980) has suggested this possibility. Theoretically, it is possible that if response classes are functionally related, the stimulus generalization of one response class may mediate the stimulus generalization of another response class. For example, suppose that a student has been trained to assemble some product that requires the use of a screwdriver and a wrench. There is natural variation in screws and bolts to which the student should generalize (generalization to screws and bolts that show variation in color is appropriate because a steel colored fastener will hold objects together as well as a brass colored fastener). Screwing and bolting response classes are potentially related because they are similar in the effects produced when used and similar in some topographical characteristics. Given the similarities in the response classes, a functional relationship (i.e., a response interrelationship) may exist between the bolting and the screwing response classes. Therefore, it is possible that programming which promotes the generalization of one response class to its corresponding stimulus class, such as bolting responses to different colored bolts, would produce the untrained generalization of the similar response class, that is, screwing responses to different colored screws. Such generalization is referred to as between-class generalization (Parsonson & Baer, 1978).

The model I tested is an extension of the strategy of "training sufficient exemplars" (Stokes & Baer, 1977). In this model, stimulus sets, in contrast to individual stimuli, are treated as exemplars of a potentially functional class. First, sets of stimuli are identified that are distinct from each other. In addition, each set must be potentially associated with a unique response class. Next, S-R relationships are established between one example of each stimulus set and its corresponding response class. Finally, training is sequentially introduced to promote generalization within some of the stimulus sets. As within-stimulus set generalization is sequentially trained across a variety of response classes, generalization probes are conducted with the remaining stimulus sets. After some sufficient amount of generalization training, spontaneous generalization of sets of stimuli may occur to their respective response classes. The model is directly analogous to the training of sufficient exemplars because new sets of stimuli can be progressively trained until spontaneous generalization occurs between other responses and untrained sets of stimuli.

The major purpose of the study was to assess the effects of within-stimulus set generalization training across related responses on the subsequent generalization of other related responses.

METHOD

Participants and Setting

Four children attending classes for moderately and severely handicapped students participated in the study. The participants' classrooms were located in a regular elementary school building and were operated by a public school system. The participants engaged in unstructured toy play with nonhandicapped children on a regularly scheduled

Table 1
Descriptive Characteristics of Students

Participant	Age (years, months)	Primary handicapping condition	IQ estimate[a]	Mean performance across subscales[b] (percentile)
Mick	7, 10	Severely handicapped, Down's syndrome	37	50th
Charles	7, 5	Severely handicapped, Down's syndrome	25	22nd
Jim	4, 2	Moderately handicapped	50	65th
Jane	4, 6	Moderately handicapped, Down's syndrome	50	70th

[a] Stanford-Binet form L-M.
[b] AAMD Adaptive Behavior Scale, TMR norms.

basis. The participants were selected because they displayed low rates of appropriate toy manipulation. Summaries of recent test results and descriptive data are given in Table 1.

Mick spoke in two-word phrases and could label a large variety of objects. Receptively, he could carry out commands such as "Turn off the lights" or "Go get a waste basket." Mick had been trained to complete many self-care skills; however, he still required instruction in zipping, buttoning, and shoe tying. He could learn new responses through imitation.

Charles rarely produced spontaneous speech, although he was capable of labeling responses. Receptively, he responded to two- or three-word commands such as "Look at me" or "Go to the door." Charles was not toilet trained and could not chew solid foods. He displayed no imitative responses during instruction.

Jim could follow two- or three-word commands. He spontaneously greeted familiar people and asked questions, (e.g., "What's that?"). His maximum utterance was four words, although he typically spoke in two-word utterances. He had been taught to identify several printed words on sight, but demonstrated inconsistent comprehension of sight words. He was capable of learning through imitation.

Jane could respond correctly to two-word commands and could label a variety of objects. She knew the names of the five other children in her class. She could produce three-word utterances, but she typically spoke one-word statements. She had excellent imitative ability.

All training and generalization sessions were conducted in an office adjoining the participants' special education classrooms. The sessions were conducted with the instructor seated across a table from the participant. All training and probe sessions were conducted individually. The instructor was a female graduate student in the severely handicapped area.

Materials

Each participant was exposed to 8 different sets of toys from the following 10 sets: animals, people, bugs, frogs, motorcycles, airplanes, boats, snakes, tanks, and spaceships. The sets of toys were constructed following procedures used in "general case programming" (Horner, Sprague, & Wilcox, 1982). Each set of toys contained five examples. The toys in each set varied in terms of size, color, and "abstractness." The range of abstractness within in each set was produced by selecting toys such that members of a set shared a small set of common configurational properties (see Table 2). The most abstract toy in each set consisted of cutout wood forms with no details other than the defining configurational elements. The other toys in each set were selected to possess the defining properties and progressively more and different details. For example, the most abstract toy airplane consisted of two Lincoln Logs crossed at right angles and attached with tape. The least abstract airplane was an accurate $\frac{1}{100}$ scale 747 jet.

The sets of toys were divided into three experimental groups. Four sets of toys were designated as generalization training sets. For example, Jane's

Table 2
Examples of Sequence of Addition of Details to Movement Toys and Characteristics of Reactive Toys

Toy set	Defining properties	Sequence of additional details
Movement		
Airplanes	Fuselage cylindrically shaped and rounded wing surfaces	1. Windows, markings, engines, wheels, surface detail, cockpit, tail 2. Windows, markings, engines, wheels, surface detail 3. Windows, markings, engines 4. Windows 5. Abstract shape, just defining properties
Boats	Rectangular section with triangular, boat-shaped front surface	1. Markings, engine, rudder, cabin, surface detail 2. Markings, engine, rudder, cabin 3. Markings, engine, rudder 4. Markings 5. Abstract shape, just defining properties
Reactive		
Windups		1. Drill 2. Bear 3. Car
Keyboard instruments		1. Small plastic piano 2. Magic flute (an electronic, plastic rod with colored keys) 3. Full-size piano

Note. Descriptions of all toy sets used in the study are available on request.

generalization training sets were snakes, boats, tanks, and people. Another four sets were designated as generalization probe sets. For example, Jane's generalization probe sets were animals, airplanes, bugs, and spaceships. Finally, two sets of toys (windups and keyboard instruments) served as an additional group of generalization probe sets. This second group of generalization probe toys was added to assess the spread of between-set generalization to toys that produced substantially different effects. That is, all other toy sets in the study were played with by physically moving the toy through some pattern of responses. In contrast, both the windup toys and the keyboard instruments produced effects that were more reactive in nature; that is, once a response is made with the object (either winding it up or pressing a key), the object itself produces an effect that is potentially noticeable. Because the reactive toys produce distinct effects from the other toys, they were analyzed separately. The sets of reactive toys contained only three objects each because multiple examples of keyboard instruments were difficult to locate. Table 2 shows the characteristics of the reactive toys as well as examples of those requiring movement responses.

For each participant, the movement-related toys were randomly designated as either generalization training or generalization probe sets. However, the assignment was controlled so that no one toy set was used more than twice in either group of toys across the four participants. In addition, if a toy was used once (or twice) in either the generalization probe or training groups it was used once (or twice) in the other group of sets.

Response Definitions

The responses to be taught were specific to each set of toys. For example, with spaceships, the participants were taught to move the toy through the air in a circular motion and land it at a right angle to the table. In contrast, airplanes took off from the table at a lesser angle and flew in straight lines. A summary of toy types and responses is given in Table 3.

Procedures

Baseline probes. The participants received a minimum of two trials with each of the 46 toys

Table 3
Response Definitions

Toy type	Response
Movement	
Airplanes	Hold plane, move plane through the air at angles less than 90°, land at angles less than 90°
Spaceships	Hold spaceship, move spaceship in circular pattern, land spaceship at 90° angle
Boats	Hold boat by its top, move on the floor, pitching nose of boat up and down
Tanks	Hold tank by its top, move slowly in a straight line, then make a sharp 90° turn
Animals	Hold animal by its top, move on the floor, move back and forth while in motion to simulate movement of limbs
People	Hold doll by back or front, move side to side during motion to simulate walking
Bugs or frogs	Hop or jump toys in a straight line
Snakes	Move toy side to side while in forward motion to produce a wavelike movement
Motorcycles	Grasp by top, move in straight line and raise front end while moving at least 6 inches to simulate a "wheelie"
Reactive	
Windups	Observe toy to find round key, rotate key until resistance is felt, place on table and observe
Keyboard instruments	Produce sequence of notes by pressing keys starting with middle key followed by the next two adjacent keys (e.g., the notes C, D, E).

to be used during the study. Verbal praise was given during the probes by saying "Good working" before the trainer showed the participant a toy. Praise was given during baseline sessions to keep the student's level of interest in the task relatively constant throughout the session and to keep the density of praise fairly constant between baseline and training trials (although this was not systematically controlled). Toys were handed to the participant with the instruction, "Play with this." The participant was then given 15 s to play with the toy.

Training with the first example toys from the generalization probe sets. Following the baseline probes, the participants were trained to produce the specific responses with the most detailed and realistic toys from each generalization probe set ("first examples"). During this training phase, the participants were also trained to play with one keyboard and one windup toy. One session was conducted each school day; each session contained 15 training trials.

The trials began with the instructor saying "Play with this." The instructor then handed the participant the toy and observed whether or not the correct sequence of responses was produced. If the student produced the correct response pattern within 10 s, enthusiastic verbal praise was delivered. If the student did not produce the correct pattern, the instructor said "No, do it like this," and simultaneously modeled the correct sequence. If the student then correctly imitated the response, the instructor said "Good" and presented the next toy to be trained. If the participant did not correctly imitate the response, the instructor said "No, do it this way." The instructor then physically guided the responses by placing the participant's hand on the toy and guiding the correct movement. No verbal praise or feedback followed manually guided responses. The criterion for ending training with a toy was set at three consecutive correct responses. Training was conducted in a spaced-trial format in that maintenance and generalization probe trials with other toys were interspersed between instructional trials. Including training, maintenance, and generalization trials, sessions typically lasted 15 min.

Generalization training with movement-related toys. After the participants reached criterion with the four first examples from the generalization probe sets, generalization training with other movement-related toy sets was begun. A multiple exemplar strategy was used to promote generalization within the training sets (Stokes & Baer, 1977).

The participants were first trained with the most detailed, realistic toy from each generalization training set. After the training criterion was met with that toy, the more abstract toys within the set were trained one-by-one until generalization to the remaining untrained toys in the set occurred. The order of introduction of the generalization training sets was randomly determined for each student. The training procedures were identical to those used during the previously described training phase. The criterion for switching from one toy set to another was either (a) when the participant generalized to all remaining toys in a set, or (b) when training was completed with all toys within a set to which the student had not generalized. Each session lasted 15 min and contained 15 training trials.

Generalization probes. The experimental sessions were organized so that probe trials were randomly interspersed among training trials. A maximum of seven toys per day was probed. The probe trials began with the statement, "Play with this," as did the training trials; however, during probe trials neither prompting nor praise was given. Generalization probes were conducted with untrained movement-related toys as well as with the untrained reactive toys.

Maintenance probes. Each of the four "first example" toys from the movement-related sets, as well as the two reactive toys that were trained during the first training phase, was probed throughout the duration of the study to ensure that the responses were maintained. If the responses were incorrect during a maintenance probe, the correct pattern of behavior was prompted as during the training trials, to ensure that the responses remained in the participant's repertoire of play responses. Correct responses were praised.

Measurement and Reliability

The dependent measure during all experimental sessions was the frequency of correct responses for each training or probe toy. A correct response was defined as producing the exact pattern of behavior defined for a given toy within 10 s of receiving the toy (see Table 3).

Totaled across the four participants, 148 sessions were conducted. Reliability probes were taken 20 times (with at least one probe under each experimental condition and with each student) by the instructor and me. Each observer independently scored the child's play as to the occurrence or nonoccurrence of the correct pattern of responses for that toy as defined in Table 3. Interobserver reliability was calculated by dividing agreements by agreements plus disagreements, times 100. Reliability was calculated on a point-by-point basis (Kazdin, 1982). The session reliability for the occurrence of target responses ranged from 82% to 100% with a median of 100%. The session reliability for nonoccurrences was 100% for all sessions except one session for which the percent agreement was 89%.

Design

A multiple probe design was used (Horner & Baer, 1978). The multiple probe data were collected within a design that conformed to a multiple baseline across responses design (Hersen & Barlow, 1976; Kazdin, 1982; Kratochwill, 1979). A multiple baseline analysis was conducted during the first training phase of the study: the training of the four "first example" toys from the generalization probe sets. A separate multiple baseline analysis was conducted with the generalization training data.

RESULTS

Training with the first example toys from the probe sets. The percentage of correct play behaviors with the most detailed toys ("first examples") from the four probe toy sets is represented in Figure 1. The baseline data across the four participants show that no correct responses were produced. Once play training was introduced, correct play responses rapidly increased with only one exception (Charles' correct responses with spaceships did not increase until the sixth training session). As training proceeded across the four toys, the participants displayed high levels of correct responses.

Generalization training. After the participants had acquired the specific responses taught with the four first example toys from the generalization probe sets, training was begun within the gener-

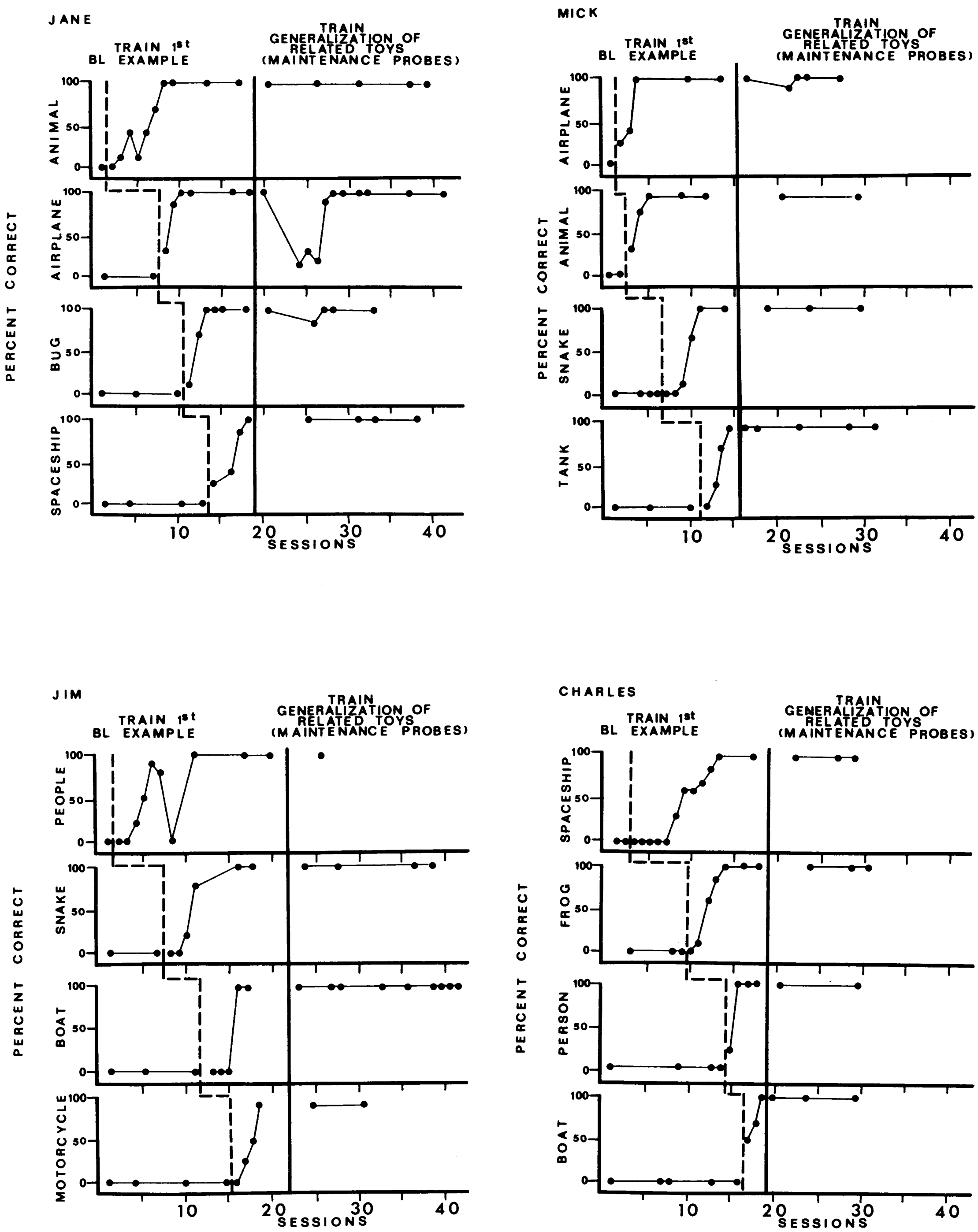

Figure 1. Percent correct play responses with the first examples from the generalization probe sets during baseline, training, and maintenance conditions.

alization training sets. In general, the first two sets required more exemplars trained than did subsequent sets. Across the four participants, the first sets required training with an average of 3.25 toys out of 4, the second set required an average of 3.5 toys trained, and the third and fourth sets required an average of 1.25 toys trained.

The generalization training data for Mick are represented in Figure 2, and they are representative of the other participants' performance. Toy boats were the first set of toys that received multiple exemplar generalization training. The data show that after training had begun with the first three exemplars, correct responses occurred with the fourth boat in the absence of training. Altogether, four out of the five boats were trained in the first set.

The second set of toys trained was spaceships. After training was begun with the first two spaceships, correct responses increased with the third untrained spaceship. The fourth and fifth spaceships required training. Mick's third set of generalization training toys was toy bugs. After training had begun with the first toy bug, correct responses were observed with the remaining untrained bugs. The fourth set of toys trained was people. As with the third set, correct responses were observed with four toys after training had begun with the first toy from the set. Altogether, Mick required training with 10 (out of a possible 20) different toys across the four sets of toys. Correct responses were produced with the remaining 10 toys without training.

Functional Control of Between-Stimulus Set Generalization By Within-Stimulus Set Generalization

Figure 3 shows the effects of generalization training across four sets of toys on the subsequent correct responses with the untrained toys from four sets to which only the first example had been trained (i.e., the generalization probe sets). Within Figure 3, the graphs that are inset to the right show the cumulative number of correct play responses to untrained toys within the generalization training sets. Each unit increase of the cumulative function represents correct responses (see Table 2) with a different untrained toy. The longer graphs underneath each inset graph show the cumulative number of correct responses with the untrained toys from the generalization probe sets. Again, correct responses (specific to a given toy set as given in Table 2) with each different toy resulted in a unit increase of the cumulative function.

Mick's data indicate that between-set generalization (i.e., correct responses to the untrained toys in sets from which only the first example was trained) did not begin until generalization training had proceeded to the second set of toys. Correct responses to untrained toys from the probe sets continued to occur as generalization training proceeded through the third and fourth sets of toys. By the end of within-set generalization training, Mick had correctly responded to all 16 of the untrained generalization probe toys.

The data for Charles are represented in Figure 3 immediately below Mick's data. The inset graph shows that Charles correctly responded to 14 untrained toys across the four sets of generalization training toys. The lower graph for Charles shows that on the last day of training within the first generalization training set, correct responses occurred with one toy from a probe set. As training progressed through the second and third sets, Charles correctly responded to progressively more toys from probe sets. By the end of training Charles had responded correctly to 9 out of 16 toys from the sets to which only the first example had been trained.

Jane's data indicate that within-stimulus set generalization training produced correct responses to 10 untrained toys across the four sets. Correct responses to untrained toys from the generalization probe sets began during generalization training within the second set of toys. As generalization training proceeded through the second, third, and fourth sets, Jane made progressively more correct responses to toys from the probe sets. By the end of training, correct responses had occurred to 13 of the 16 untrained toys from the generalization probe sets.

Jim's data indicate that correct responses were produced with three toys prior to the initiation of generalization training. These correct responses occurred with three examples from the set of toy people. During generalization training with the first two sets of toys, Jim made fewer correct responses to untrained toys than did the other participants. However, Jim did correctly respond to the maximum possible number of untrained toys within the third and fourth generalization training sets. Thus, with the exception of the toy people, Jim followed

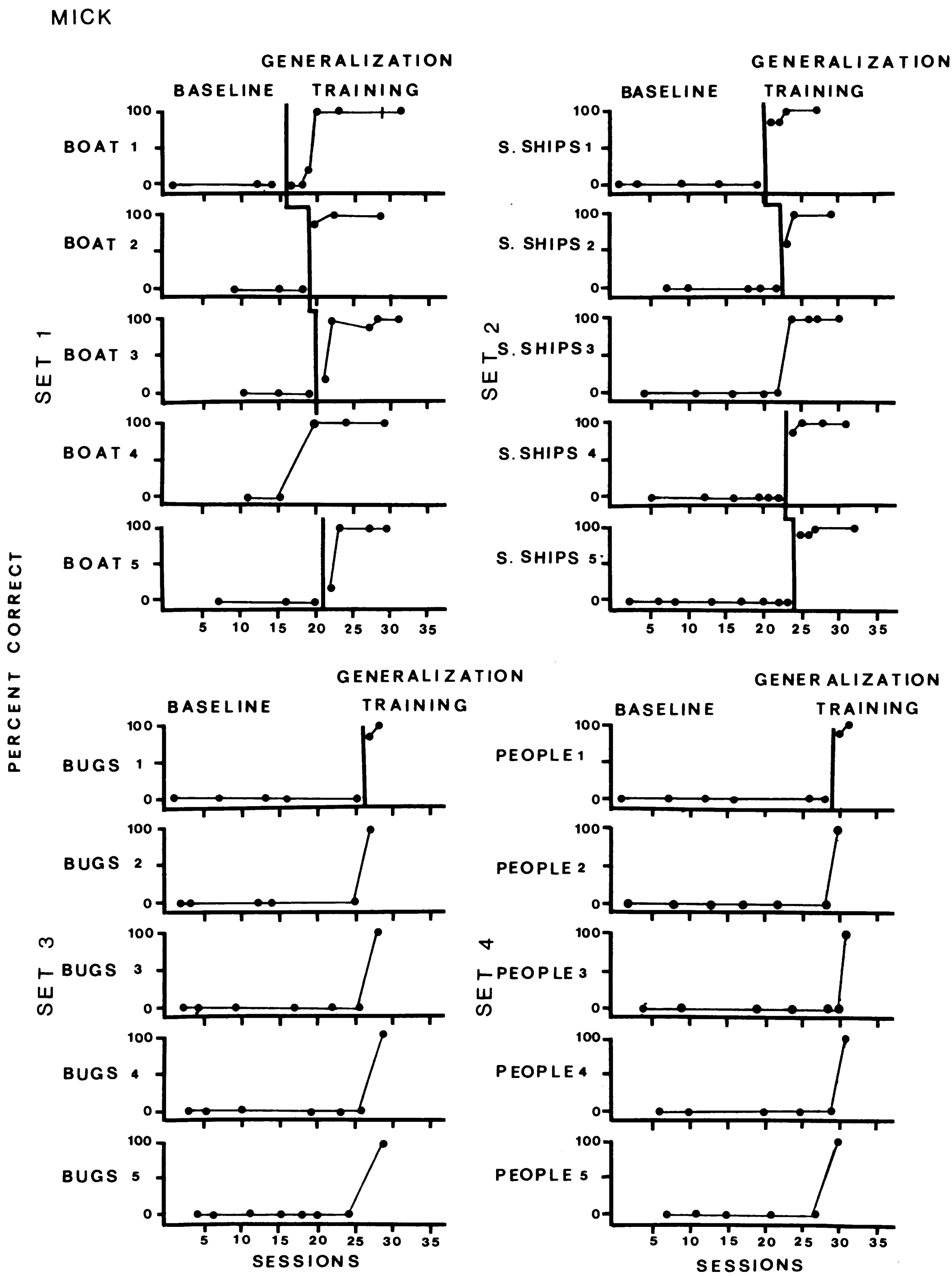

Figure 2. Results for Mick of within-stimulus set generalization training across four sets of toys.

a similar pattern to the other participants in that correct responses to toys from the probe sets did not occur until within-stimulus set generalization training had progressed to the second set. Altogether, Jim made correct responses to seven untrained toys from the probe sets (11 if people are included as they are in Figure 3).

In summary, generalization training within two

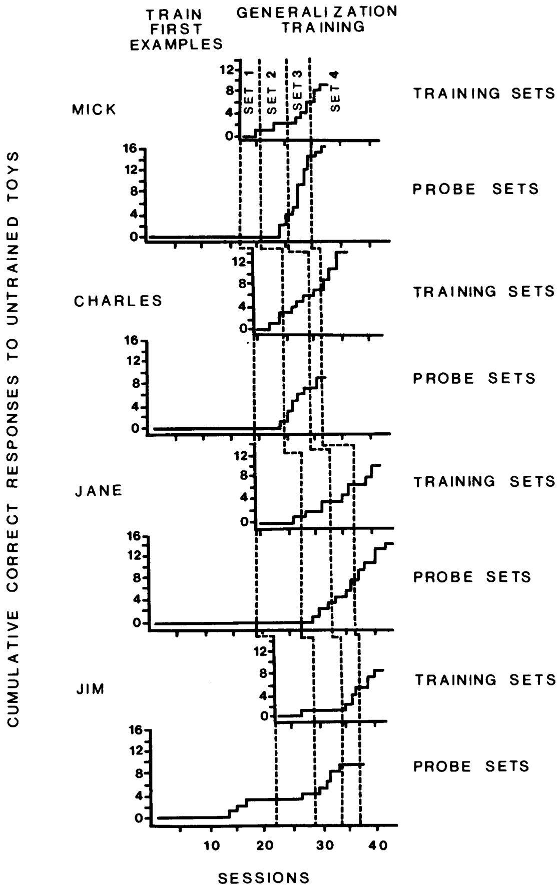

Figure 3. Cumulative correct responses to untrained toys across the four participants. On the inset upper graph for each participant, the cumulative correct responses to untrained toys from generalization training sets is shown. On the lower graph for each participant, unreinforced probes for between-stimulus set generalization during baseline, first exemplar training, and within-stimulus set generalization training is shown.

sets was associated with the beginning of correct responses to untrained toys from the probe sets for Mick and for Jane. Although Charles began to respond correctly to probe set toys during the first generalization training set, his maximum rate of between-stimulus set generalization occurred during the second generalization training set. Jim's data indicate that generalization to one set of probe toys (people) occurred prior to within-set generalization training; however, as with Mick and Jane, correct responses to untrained toys occurred most frequently during training within the second generalization training set.

Patterns of Between-Stimulus Set Generalization

Figure 4 shows the percentage of toys from each generalization probe set that were correctly responded to by the participants. Ten sets of movement-related toys were used; six of the sets were used with two students (people, airplanes, animals, snakes, motorcycles, and boats), and the remaining four sets were used with only one student. The participants responded correctly with all the toys from 7 out of 10 of the movement toy sets. Within the set of motorcycles, Jim responded correctly to three out of four toys. Within the set of boats, one out of the two students who received the set did not respond correctly to any of the toys in the set, and the other student responded to only one boat. The same results were observed with spaceships. In contrast to the movement toys, no correct responses were observed with any of the reactive toys.

DISCUSSION

In general, the results indicate that when generalization training had proceeded within one or two sets of toys, there was a facilitation of generalization (between toy sets) to other untrained toys (see Figure 3). Although the degree of generalization observed was impressive, there was little or no generalization to two sets of movement-related toys (boats and spaceships). With the exception of one of Jim's toy sets, the occurrence of between-class generalization was dependent on exposure to within-class generalization training.

The results indicate that Jim generalized to three out of four toy people prior to the beginning of generalization training. It should be recalled that the first toy that Jim was trained with was the first example of toy people (the set included a small "Star Wars" android figure, a "Troll" doll, a "Gumby," a male doll dressed in conventional clothing, and a cutout wood figure with no detail). After Jim had been trained to play with the first toy person, he began to produce the people response class with other toys in the study. With the exception of the full-size piano, at one time or another Jim produced the people response with

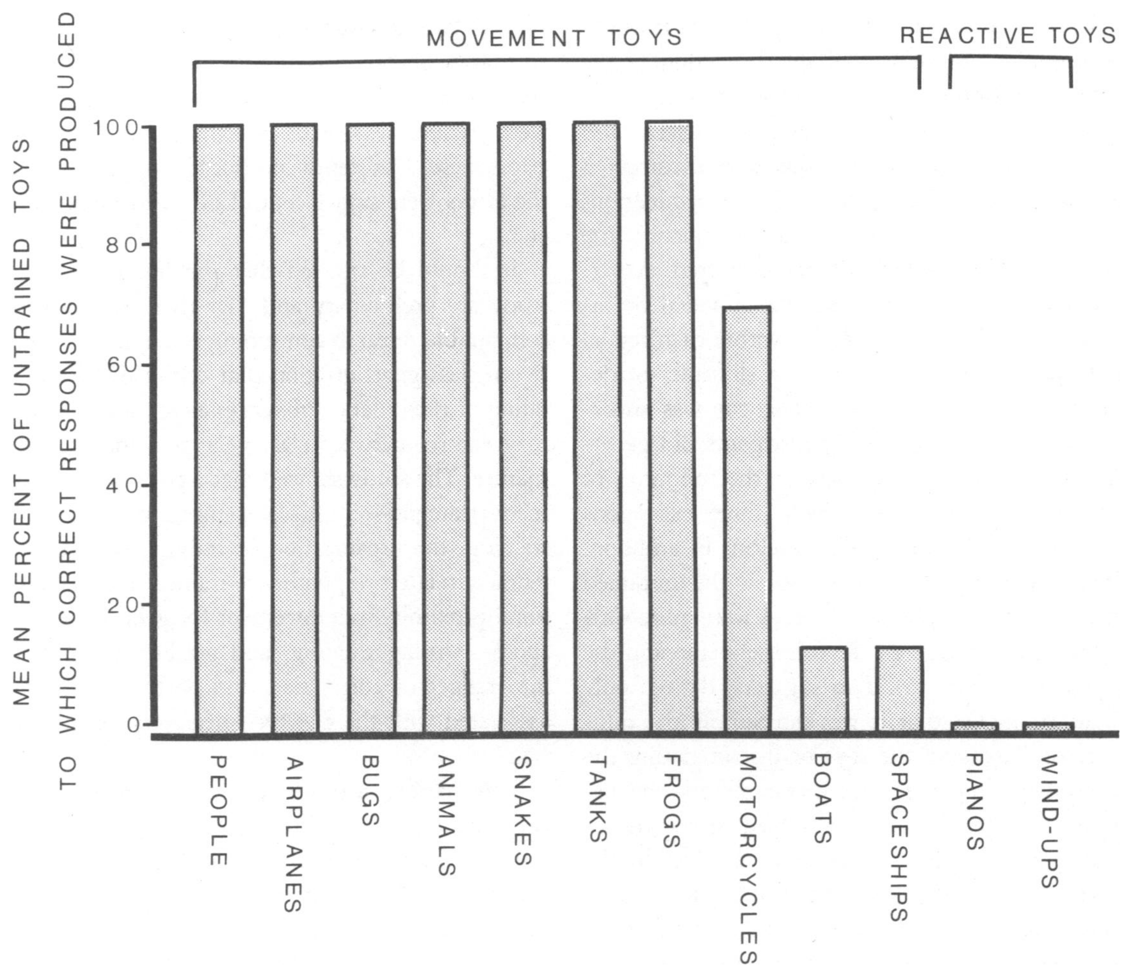

Figure 4. Mean percentage of untrained toys to which correct responses were made during generalization probes with movement toys and reactive toys.

every toy in the study. Thus, although Jim's data indicate that generalization occurred prior to the introduction of generalization training, those responses probably represent a form of generalization that is based on a failure to discriminate differences between toys.

Importantly, generalization was not observed to the toys from the sets of reactive toys. This failure to generalize could be due to several possible factors: the discrimination of the defining properties of the reactive toys may have been more difficult, the students may have had fewer real life experiences or histories of play with toys similar to those from the reactive sets, or some critical relationship to the other toy sets may have been lacking. The characteristics of stimuli or responses that control the spread of between-stimulus set generalization warrant further discussion and experimentation. In this investigation, the movement-related toys required similar response topographies (e.g., holding the toys and moving them in similar patterns). Thus, it is possible that similarities in response topographies control between-class generalization. It is also possible that similarities in the stimulus features that require discrimination may control between-stimulus class generalization. Finally, similarities or differences in effects may exert control. In this study, the reactive toys differed from the movement-related toys on at least two of these dimensions—topographies and effects.

For a clearer interpretation of these data it would be important to show that discrimination of the defining attributes of each set was of comparable difficulty across the sets of movement-related and reactive toys. Although the sets were constructed so that at a subjective level the discriminations required seemed to be of comparable difficulty, the similarity is not empirically demonstrated. A partial control for this problem was provided by including a wide range of objects within each set, so as to produce a realistic range in difficulty of determining whether or not a given toy was an example of a set. When the participants did generalize to a set, they generalized to the full range of toys within the set with only three exceptions (spaceships, boats, and motorcycles). In addition, the participants did not generalize to the untrained reactive toys even though the toys were quite similar in some cases (e.g., the full-size piano and the smaller plastic piano). This suggests, though only circumstantially, that it was not simply the difficulty of classifying the toys or discriminating the controlling properties that accounted for the between-class generalization observed to the movement toys and the lack of generalization to the reactive toys. If this argument can be made more convincing (with additional studies in the future), these data may indicate that if generalization is an operant that can be trained as Parsonson and Baer (1978) suggested, the variables controlling a generalized operant may be relatively specific to the task, materials, or context within which the responses were trained and probed.

The lower degree of between-class generalization to boats and spaceships is problematic because those sets were organized in the same manner as the other movement sets. The errors in generalizing to spaceships were due to substituting the airplane response class for the spaceship class. The errors produced with the toy boats usually involved moving the boat back and forth without simulating a wave pattern. The boat responses that were scored as errors were partially correct. Interestingly, the error responses produced with spaceships and boats were based on consistent patterns (i.e., substitutions or partial omissions) rather than random substitutions of other responses. Although the responses with toys from the spaceship and boat sets were frequently scored as error responses, those errors were consistently produced, suggesting that the errors were under control of a generalized operant.

It should be stressed that the findings of this study are preliminary and that there is a lack of comparable research concerning the training of between-class generalization that could aid interpretation of these data. Inferences as to variables that may control between-class generalization are premature. The study showed that a package of treatment strategies—multiple exemplar generalization training, the organization of training so that potentially related responses are trained in close temporal proximity, reinforcement for generalizing responses during training, and grading the objects into ranges of color, size, and abstractness—was associated with the observed degree of generalization.

Explanations accounting for the formation of response-response relationships could be either: (a) the close temporal occurrence of responses, (b) the functional similarity of the responses in producing some effect, (c) similar antecedent, controlling variables, or (d) some combination of these. It may be useful to investigate the formation of response interrelationships (as evidenced by between-class generalization) with a finer grained analysis to identify stimulus- and response-related features that may control generalization.

In conclusion, this study demonstrated a training strategy based on the theoretical influence of response interrelationships on stimulus generalization. It is apparent that there are a number of ways in which responses can form interrelationships, and there are multiple effects that such relationships may exert on the learning, performance, and generalization of responses. It is hoped that continued research in this relatively new area of investigation will lead to increased efficiency of teaching programs without concomitant increases in the complexity of teaching technology.

REFERENCES

Casalta, C. H. (1980). Generalizacion mediada: La construccion de interacciones complejos entre respuestas. *Revista Mexicana de Analisis de la Conducta*, **6**, 167–183. (From *Psychological Abstracts*, 1982, *68*, Abstract No. 7341.)

Guess, D., & Baer, D. M. (1973). An analysis of individual differences in generalization between receptive and productive language. *Journal of Applied Behavior Analysis*, **6**, 311–329.

Haring, T. G., Breen, G., Pitts-Conway, V., & Gaylord-Ross, R. (1984). *The use of differential reinforcement of other behaviors to reduce stereotyped behavior of autistic students during group instruction*. Manuscript submitted for publication.

Hersen, M., & Barlow, D. H. (1976). *Single-case experiment designs: Strategies for studying behavior change*. New York: Pergamon.

Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, **11**, 189–196.

Horner, R. H., Sprague, J., & Wilcox, B. (1982). General case programming for community activities. In B. Wilcox, & G. T. Bellamy (Eds.), *Design of high school programs for severely handicapped students* (pp. 61–98). Baltimore: Paul H. Brooks.

Husaim, J. S., & Cohen, L. B. (1981). Infant learning of ill defined categories. *Merrill-Palmer Quarterly*, **27**, 443–456.

Kara, A., & Wahler, R. G. (1977). Organizational features of a young child's behaviors. *Journal of Experimental Child Psychology*, **24**, 24–39.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. Oxford: University Park.

Koegel, R. L., & Covert, A. (1972). The relationship of self-stimulation and learning in autistic children. *Journal of Applied Behavior Analysis*, **5**, 381–388.

Kratochwill, T. R. (1979). *Single-subject research: Strategies for evaluating change*. New York: Academic.

Lee, V. L. (1981). Prepositional phrase spoken and heard. *Journal of the Experimental Analysis of Behavior*, **35**, 227–242.

Lichstein, K. L., & Wahler, R. G. (1976). The ecological assessment of an autistic child. *Journal of Abnormal Child Psychology*, **4**, 31–54.

Neef, N. A., Shafer, M. S., Egel, A. L., Cataldo, M. F., & Parrish, J. M. (1983). The class specific effects of compliance training with do and don't requests: Analogue analysis and classroom application. *Journal of Applied Behavior Analysis*, **16**, 81–99.

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

Piaget, J. (1980). *Adaptation and intelligence: Organic selection and phenocoy*. Chicago: University of Chicago Press.

Russo, D. C., Cataldo, M. F., & Cushing, P. J. (1981). Compliance training and behavioral covariation in the treatment of multiple behavior problems. *Journal of Applied Behavior Analysis*, **14**, 209–223.

Sherman, J. A. (1964). Modification of nonverbal behavior through reinforcement of related verbal behavior. *Child Development*, **35**, 717–723.

Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *The Journal of General Psychology*, **12**, 40–65.

Skinner, B. F. (1953). *Contingencies of reinforcement: A theoretical analysis*. New York: Macmillan.

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

Strain, P. S., & Ezzell, D. (1978). The sequence and distribution of behavioral disordered adolescents disruptive/inappropriate behaviors: An observational study in a residential setting. *Behavior Modification*, **2**, 403–425.

Voeltz, L. M., & Evans, I. M. (1982). The assessment of behavioral interrelationships in child behavior therapy. *Behavioral Assessment*, **4**, 131–165.

Wahler, R. G., Sperling, K. A., Thomas, M. R., Teeter, N. C., & Luper, H. L. (1970). The modification of childhood stuttering: Some response-response relationships. *Journal of Experimental Child Psychology*, **9**, 411–428.

Whitehurst, G. J. (1977). Changes in verbal behavior collateral to nonverbal training. *Journal of Experimental Child Psychology*, **23**, 23–28.

Received December 12, 1983
Final acceptance March 5, 1985