JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1994, 27, 371–384 NUMBER 2 (SUMMER 1994)

ASSESSING AND PROGRAMMING GENERALIZED BEHAVIORAL REDUCTION ACROSS MULTIPLE STIMULUS PARAMETERS

BRIDGET A. SHORE, BRIAN A. IWATA, DOROTHEA C. LERMAN, AND MELISSA J. SHIRLEY

THE UNIVERSITY OF FLORIDA

Generalization across three stimulus parameters was examined for 5 individuals whose self-injurious behavior was maintained by escape from task demands. Prior to treatment, three stimulus parameters (therapist, setting, and demands) were systematically varied across baseline sessions. These variables were held constant during treatment, which consisted of escape extinction. When treatment was completed, three novel stimulus parameters were probed. If the rate of self-injury was high during this probe, treatment was reimplemented with one new stimulus parameter (the other two were the same as in the original treatment condition). Following this second treatment, another probe with three novel stimuli was conducted. If the rate of self-injury was again high, treatment was implemented again while a second stimulus parameter was changed. This sequence continued until generalization was observed across the three parameters. Results showed idiosyncratic differences in generalization. The behavior of 2 subjects showed complete generalization during the first novel probe. A 3rd subject's behavior showed generalization following treatment across two stimulus parameters (setting and therapist). The behavior of the 2 remaining subjects showed a complete lack of generalization across the three parameters; both subjects required training for novelty by randomly varying the stimulus parameters for a substantial number of sessions.

DESCRIPTORS: extinction, functional analysis, generalized behavioral reduction, self-injurious behavior, stimulus generalization

Baer, Wolf, and Risley (1968) proposed seven guidelines for the conduct of applied behavior analysis research, the last of which was that behavioral procedures should be capable of generalized outcomes. They emphasized the importance of explicitly examining generalization, that generalization should not be assumed to occur automatically, and that it should be programmed rather than expected or lamented. Since that time, a considerable amount of research has examined generalized behavior change across settings, responses, experimenters, and/or time subsequent to initial intervention (Baer, Wolf, & Risley, 1987).

Stokes and Baer (1977), in reviewing the methodology of 270 published studies, identified a number of distinct strategies for assessing or programming generalization. Their analysis was intended to encourage and direct future research examining the variables that produce generalization. Their taxonomy proved to be fruitful, in that a number of subsequent studies examined variations of the techniques they described. For example, perhaps the most frequently used technique is "train sufficient exemplars," which involves teaching new exemplars one after another until induction is formed. Generalization is thus programmed by training a subset of the stimulus conditions or responses targeted for generalization. Most research in this area has dealt with generalization across experimenters (e.g., Corte, Wolf, & Locke, 1971; Stokes, Baer, & Jackson, 1974), although research has also examined generalization across settings (Corte et al., 1971; Rincover & Koegel, 1975) and across response topographies (Garcia, Baer, & Firestone, 1971). Recent variations of this technique have been used to promote generalization across conversational topics and partners during speech training for patients with Broca's aphasia (Doyle, Gold-

This research was supported in part by a grant from the Florida Department of Health and Rehabilitative Services. We thank Iser DeLeon, Beth Duncan, Han-Leong Goh, Jodi Mazaleski, Joel Ringdahl, Richard Smith, and Jennifer Zarcone for their assistance in conducting various aspects of the study.

Reprints may be obtained from Brian Iwata, Psychology Department, The University of Florida, Gainesville, Florida 32611.

stein, Bourgeois, & Nakles, 1989) and to promote generalized compliance with requests made by novel adults (Davis, Brady, Williams, & Hamilton, 1992).

Kirby and Bickel (1988) extended Stokes and Baer's (1977) analysis by discussing the behavioral principles underlying generalization. They identified stimulus control as the mechanism that produced stimulus generalization and presented an analysis of generalized behavioral effects based on this assumption. This analysis emphasized three important points concerning the establishment of stimulus control as it relates to generalization: (a) Conditional controlling relations between stimuli and behavior may prevent the stimuli that acquire control over responding from exerting their control in all contexts, (b) unintended correlations (not explicitly defined by the experimenter) may acquire control over behavior, and (c) apparently simple stimuli consist of dimensions or components that may exert control either separately or in combination. Therefore, successful generalization programming requires analysis and planning to establish appropriate stimulus control, and, when behavioral effects fail to generalize, further analysis is needed to identify stimulus control relations that had been overlooked.

Although considerable research has examined generalization per se, systematic assessment of the multiple stimulus parameters that may affect generalized behavioral outcomes has rarely been conducted. One recent study that did attempt this type of analysis was conducted by Halle and Holt (1991). Four moderately mentally retarded students in a vocational training school were taught to use the word "please" in the context of running errands for the teacher. Following training in which four stimulus parameters remained constant (requester, item, receiver, and setting), probes were interspersed with training sessions to assess generalization of the response "please" to conditions with novel stimulus parameters. Halle and Holt demonstrated that training subjects to respond in the presence of invariant stimulus parameters provided a method for the direct manipulation of those parameters after training to assess which stimuli had become discriminative for the newly acquired response.

The assessment of generalization following treatment to reduce the frequency of inappropriate behavior presents a somewhat different case than the assessment of generalization subsequent to acquisition training. When a new skill is learned, one can look for stimulus control in the absence of contrived consequences, because, in most cases, contingencies in the natural environment should be expected to maintain the behavior. The continued reduction of inappropriate behavior, however, requires that maintaining environmental consequences remain altered. Therefore, when examining stimulus control and generalization of a newly reduced behavior, the reemergence of inappropriate behavior when stimulus parameters are varied should be examined in the absence of the reinforcing environmental consequence. Because extinction should remain in effect during generalization probes, additional problems in assessing stimulus control are encountered. First, one must limit the number of probes to prevent treatment (ongoing exposure to extinction) during novel conditions. Second, when a novel set of stimuli has been probed, it is no longer novel. Any problem behavior that occurs during the probe sessions will have contacted the consequences (or lack thereof) and therefore received treatment in the novel condition.

Escape-maintained self-injury often occurs in situations that contain at least three very salient stimulus parameters: therapist, setting, and instructional content. Thus, by using invariant stimulus parameters during treatment, probes using novel stimuli may be conducted posttreatment to assess generalization. Sequentially treating each novel stimulus parameter when generalization does not occur provides a method for systematic analysis and treatment for generalization. The purpose of this study was to assess and program generalization across several stimulus parameters subsequent to treatment for self-injury (SIB) maintained by escape from demands.

METHOD

Subjects and Setting

Five adult males living in a state residential facility for persons with developmental disabilities participated as subjects. All 5 subjects were diagnosed with severe/profound mental retardation. Michael was 44 years old, was ambulatory, had no expressive language, and seldom complied with instructions. His SIB consisted of hand biting, which was frequently accompanied by crying, flopping on the floor, and running away. Charles was 45 years old, was ambulatory, and also displayed no expressive language. He was extremely noncompliant, and his SIB consisted of head and body hitting, arm and hand biting, and face and body scratching. His SIB was also frequently accompanied by aggression and attempts to overturn furniture or destroy property. Jeff was 44 years old, displayed no expressive language, and required physical assistance to walk. He spent most of his time in a wheelchair and often wore a helmet due to his extreme instability. His SIB consisted of head banging and arm and hand biting. Jasper was 40 years old, was ambulatory, and, although he could follow many complex instructions, his expressive language was limited to infrequent one- or two-word phrases. His SIB consisted of head and body hitting and head banging. Jacob was 44 years old, was ambulatory, and, although he displayed no expressive language, he was able to follow simple instructions. His SIB consisted of hand biting and head hitting.

All sessions were conducted at a day treatment center located on the grounds of the facility. Up to six different treatment rooms at the center were used for treatment or probe sessions.

Response Measurement and Reliability

The primary dependent variables were rate of SIB and percentage compliance with instructions (demands). Self-injurious behaviors were defined as follows: ***head or body hitting***—audible contact of a hand, fist, or knee against any part of the face, head, or body; ***head banging***—contact of the head with a stationary object; ***hand or arm biting***—closure of the teeth on any part of the skin from shoulder to fingertips; and *scratching*—a raking motion of the fingertips across any part of the body. Compliance was defined as performing a requested task without physical assistance. During each session, an observer recorded each occurrence of SIB, each demand, and each compliance on a hand-held computer (Assistant Model AST 102). Rate of SIB was calculated by dividing the number of responses by session time. Percentage compliance was obtained by dividing the number of compliances by the number of instructional trials.

Interobserver agreement was assessed by having a second observer simultaneously but independently record data. Session time was divided into consecutive 10-s intervals, and agreement percentages were calculated based on interval-by-interval comparisons of the observers' records. The smaller number of responses in each interval was divided by the larger number of responses. These fractions were then summed across all intervals and divided by the total number of intervals in the session to get the percentage agreement between the two observers. The percentages of sessions during baseline and treatment for which reliability was assessed, and the mean agreement percentages for subject SIB, therapist demands, and subject compliance, respectively, were as follows: Michael (43.2% of sessions), M = 97.4%, 93.5%, and 99.9%; Jeff (34% of sessions), M = 97.7%, 96.3%, and 95.4%; Charles (25% of sessions), M = 98.9%, 96.7%, and 99.7%; Jasper (30% of sessions), M = 98.4%, 96.8%, and 97.6%; Jacob (27% of sessions), M = 97.1%, 99.5%, and 95.1%.

Functional Analysis Assessment

Before the study, an assessment was conducted to identify the variables maintaining subjects' SIB. All subjects were exposed to a series of 15-min sessions representing four conditions (attention, demand, alone, and play) that were presented in a multielement design as described by Iwata, Dorsey, Slifer, Bauman, and Richman (1982). Length of assessment for individuals ranged from 12 to 46

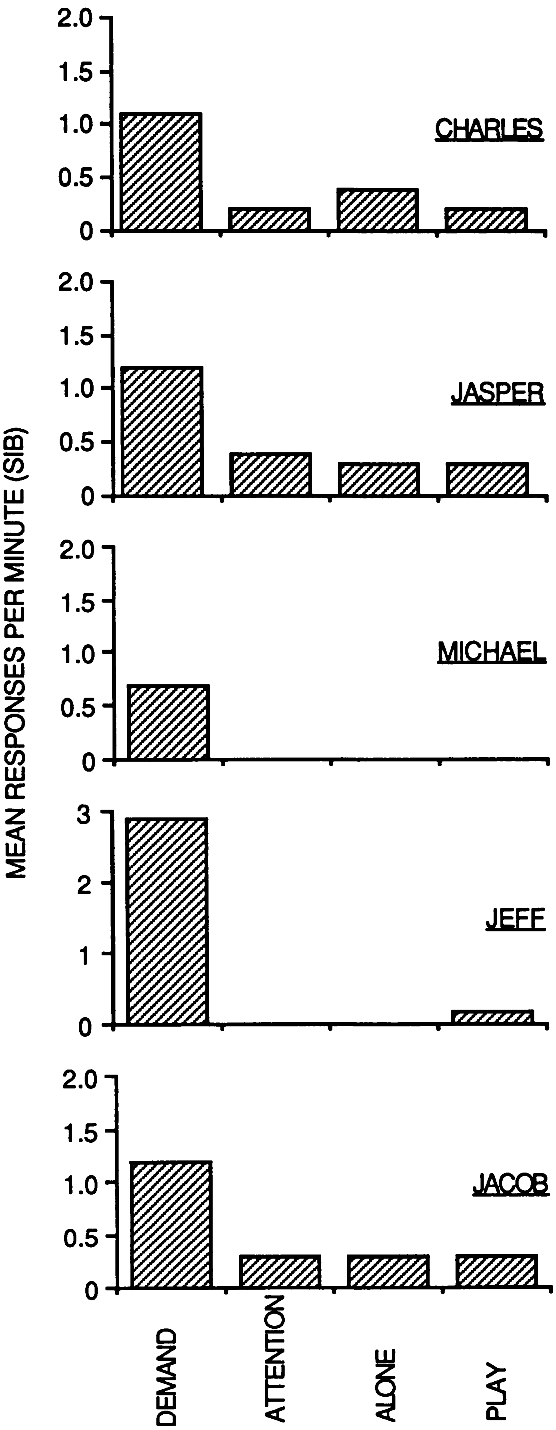

Figure 1. Mean rates of SIB across assessment conditions.

sessions (M = 19). The percentage of sessions during which interobserver agreement was assessed for individuals ranged from 36.1% to 74.2% (M = 45.8%), and mean percentage agreement scores ranged from 98.2% to 99.8% (M = 99.1%). Figure 1 shows subjects' mean rates of SIB during the assessment conditions. All subjects exhibited much higher rates of SIB during the demand condition, indicating that their SIB was maintained by escape (negative reinforcement).

Experimental Design

Three stimulus parameters were manipulated during the different conditions of the study: Up to six settings (S1 through S6), six therapists (T1 through T6), and six sets of task demands (D1 through D6) were selected for each subject prior to treatment. The specific parameters in effect for each session were also recorded (e.g., S2T2D1).

Baseline probes consisted of two sets of the three stimulus parameters varied in all eight possible combinations, presented in different order across subjects. Following the implementation of treatment with a constant set of stimuli (S1T1D1), additional probes were conducted with novel stimulus parameters, and further treatment was undertaken as necessary. Experimental control was demonstrated through replication across subjects and via reversals when periodic novel probes were interspersed with treatment conditions. The probe procedure constituted a reversal condition by reinstating novel stimulus conditions. To the extent that SIB increased in some subjects during some novel probes, a reversal effect was shown. Also, subjects usually experienced differing numbers of treatment sessions between probes, which, in addition to the reversal feature of the design, comprised a multiple baseline across stimulus parameters (analogous to "settings") for those subjects who required treatment on a given parameter.

Stimulus Parameters

Settings. Six different rooms at the treatment site were used. Briefly, the rooms were (a) a large room with windows on three sides and a wall and doorway on the third wall; (b) a smaller room with windows on two sides, one solid wall, and a large doorway into an adjacent room; (c) an even smaller room with one window, three solid walls, and a door; (d) a small entryway with a door to the outside, a door into an adjacent room, and two solid walls; (e) a medium-sized rectangular room with windows along one wall, a large table, and two doorways into adjacent rooms; and (f) a kitchen with a doorway into an adjacent room and a door to the outside. All rooms were relatively barren (e.g., no wall decorations) except for a table and two chairs (if required for some tasks) and any materials needed for task completion.

Therapists. All six therapists were graduate research assistants experienced in the treatment of SIB. None of the five novel therapists (T2 through T6) had a prior history of working with the subjects.

Demand sets. A pool of instructions was obtained by reviewing each subject's individual habilitation plan and current training programs and by interviewing staff. Each subject was then exposed to a series of instructional trials during an assessment conducted before the study. The demands were presented approximately every 30 s using a three-prompt sequence consisting of a verbal instruction, a touch prompt, and physical guidance. Compliance was defined as the subject's completing the task requirement without physical assistance, and the consequence for compliance was verbal praise. The consequence for SIB was removal of the task demand (escape). The criterion for including a given demand in the study was at least 10 trials with 35% or more trials resulting in SIB and 25% or fewer trials resulting in compliance. Up to six sets of demands with three demands each were obtained for each subject (a list of the demands used for each subject is available from the authors).

Table 1
Order of Baseline Probes for Charles, Jasper, Michael, Jeff, and Jacob

Charles	Jasper	Michael	Jeff	Jacob
S1T1D1	S1T1D1	S1T1D1	S1T2D1	S1T2D1
S2T1D2	S1T1D2	S2T1D1	S1T2D2	S2T2D2
S2T1D1	S2T1D1	S1T2D1	S2T2D1	S1T2D2
S1T2D2	S2T2D2	S1T1D2	S2T2D2	S2T2D1
S1T2D1	S2T1D2	S2T2D1	S2T1D1	S1T1D1
S1T1D2	S1T2D1	S2T2D2	S2T1D2	S2T1D2
S2T2D2	S2T2D1	S1T2D2	S1T1D2	S2T1D1
S2T2D1	S1T2D2	S2T1D2	S1T1D1	S1T1D2

Procedures

Baseline probes. Before treatment, probes were conducted in which task demands were spaced 30 s apart, for a total of 30 demands per 15-min session. As in the demand assessment, a three-prompt sequence was used, and compliance was defined as the subject's completing the task without physical assistance. The consequence for compliance was verbal praise, and the consequence for SIB was removal of the task demand (escape). Sets 1 and 2 of the stimulus parameters were varied in all eight possible combinations (S1T1D1, S2T1D1, S1T2D1, S2T2D1, S1T2D1, S1T1D2, S1T2D2, S2T2D2). Table 1 shows the order of baseline probes for each subject.

Treatment. Task demands were presented as in baseline. The prompting sequence and praise for compliance remained in effect. During treatment, however, an extinction procedure was implemented for escape behavior. SIB no longer produced termination of the demand trial; instead, the therapist continued the session as scheduled, guiding the subject through the task (Iwata, Pace, Kalsher, Cowdery, & Cataldo, 1990). For all subjects, treatment was initially implemented with setting, therapist, and demand set held constant (S1T1D1), and was considered complete when the rate of SIB was less than 0.5 responses per minute for five consecutive sessions. Once treatment in the S1T1D1 condition was completed, escape extinction remained in effect for all subsequent treatment conditions and novel probes.

Generalization probe sequence. Figure 2 shows the probe sequence used for all subjects in the study. Following the successful reduction of SIB during the first treatment condition (S1T1D1) with escape extinction, a probe session was conducted with all novel parameters (S2T2D2). If the rate of SIB was below 0.5 responses per minute for three consecutive novel probes, generalization was considered to have occurred. However, if the rate of SIB ex-

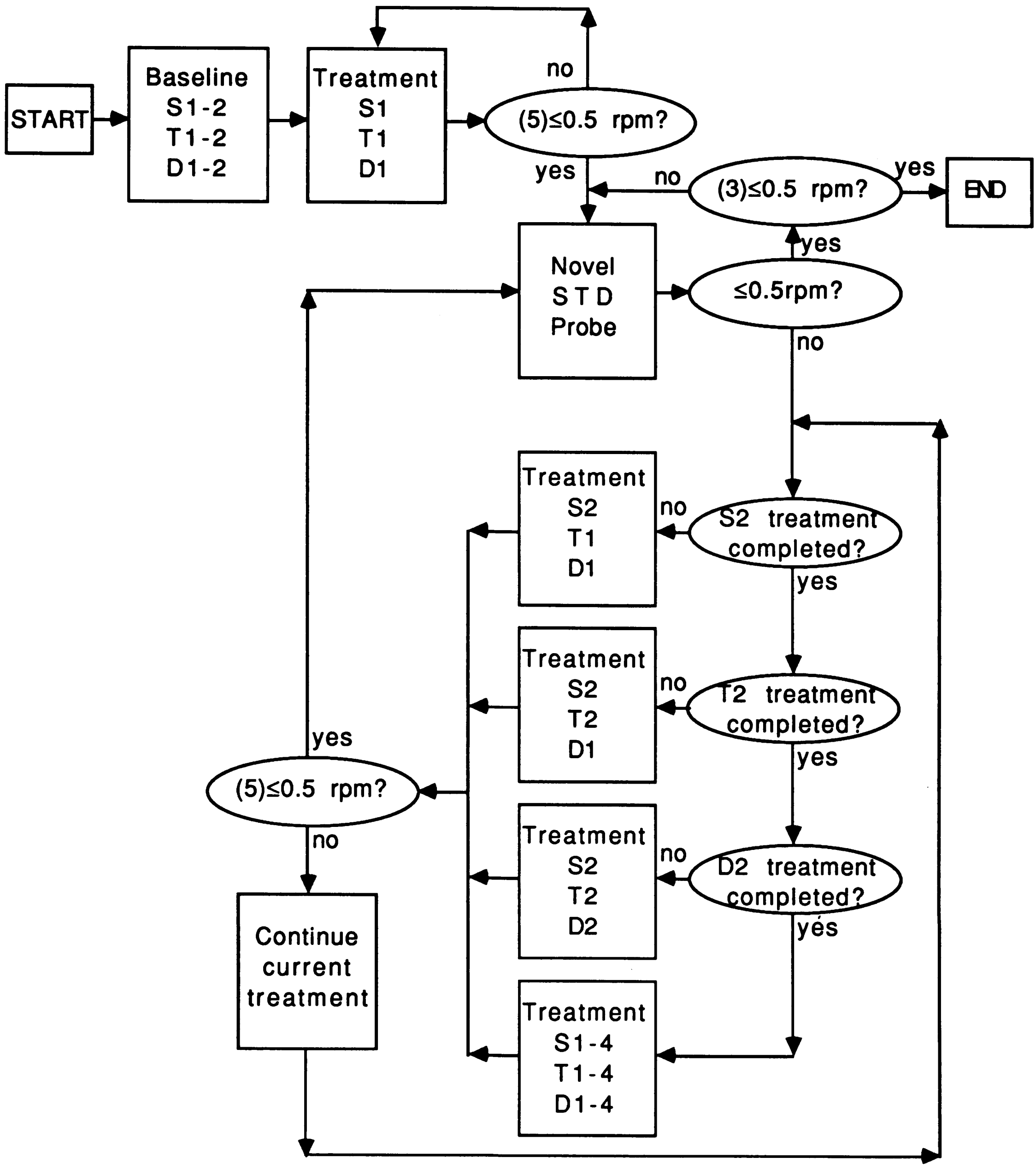

Figure 2. Flow-chart illustration of the sequence of conditions for all subjects. Letter and number codes refer to stimulus parameters consisting of setting (S), therapist (T), and demand set (D).

ceeded 0.5 responses per minute, treatment began for one novel stimulus parameter (setting), while the other two parameters remained the same as in the original treatment (S2T1D1). When the rate of SIB remained below 0.5 responses per minute for five consecutive sessions, another novel probe occurred (S3T3D3). If this probe showed SIB greater than 0.5 responses per minute, then treatment was implemented on the second stimulus parameter (therapist), with the other two parameters remaining the same as in the previous treatment condition (S2T2D1). Following successful treatment in this condition, another novel probe occurred (S4T4D4). If the rate of SIB exceeded 0.5 responses per minute, treatment was implemented on the third stimulus parameter (demand set), with the other two parameters remaining the same as in the previous treatment condition (S2T2D2). When the treatment criterion of five consecutive sessions with SIB below 0.5 responses per minute was reached, another novel stimulus probe was conducted (S5T5D5). If generalization to novel stimulus parameters (three consecutive novel probes with SIB below 0.5 responses per minute) had not occurred when all three stimulus parameters had been treated, "training for novelty" was implemented. This procedure involved varying the three stimulus parameters (e.g., S1 through S4, T1 through T4, D1 through D4) in random combinations each session until the treatment criterion of five consecutive sessions with SIB below 0.5 responses per minute was reached. Novel stimulus parameters (S5T5D5 or S6T6D6) were then probed for generalization. The order of treatment across the novel stimulus parameters (e.g., setting, therapist, and then demand set) was arbitrarily chosen but was based on the assumption that treatment in a new setting (with the same therapist and task demands) was least likely to produce generalization across all novel stimulus parameters, whereas treatment with a either a new therapist or a new set of task demands might be more likely to produce generalization.

RESULTS

Self-Injurious Behavior

Figure 3 shows the rate of SIB across conditions for all 5 subjects. During treatment conditions, if a subject required more than 10 sessions to meet criterion, data are shown only for the last 10 sessions. However, the number of sessions in each treatment condition is noted in parentheses.

Charles's baseline data showed relatively stable rates of SIB, with a mean of 0.86 responses per minute (range, 0.67 to 1.47). During the last 10 sessions of treatment in the S1T1D1 condition, his mean rate of SIB was 0.47 responses per minute (range, 0.13 to 1.6). Of the 5 subjects, Charles's initial treatment condition was the lengthiest (32 sessions), yet he showed immediate generalization to a novel setting, therapist, and demand set during the S2T2D2 probe.

Jasper's data showed considerable variability during baseline, with a mean SIB rate of 1.6 responses per minute (range, 0.4 to 5.13). Examination of the order of stimulus conditions (Table 1) did not show consistent differences among settings, therapists, or demand sets. For example, although the second session (S1T1D2) showed the highest rate of SIB (5.3 responses per minute), the third and fourth sessions also showed relatively high rates (2.2 responses per minute for S2T1D1, and 2.8 responses per minute for S2T2D2). Although there were no differences between sets, Jasper's baseline showed a downward trend across sessions. His mean rate of SIB during the last 10 sessions of escape extinction was 0.94 responses per minute (range, 0 to 3.87). He required 21 sessions of treatment to meet treatment criterion in S1T1D1, and then treatment effects immediately generalized to novel stimulus parameters (S2T2D2). He displayed SIB once during one session of the generalization probes.

Michael's baseline showed stable rates of SIB (M = 0.91 responses per minute; range, 0.53 to 1.27) across all eight combinations of stimulus parameters. He required 20 sessions to meet criterion in the first treatment condition (S1T1D1). His mean rate of SIB during this condition was 0.2 responses per minute (range, 0 to 1.4). During Michael's first novel probe (S2T2D2), his rate of SIB exceeded the treatment criterion (0.67 responses per minute). Treatment was therefore resumed, with the therapist and demand set remaining the same as during his first treatment condition, but in a new setting (S2T1D1). This treatment condition took 10 sessions to complete (mean rate of SIB = 0.74 responses per minute; range, 0 to 2.2). Another novel probe was then conducted (S3T3D3) and, although he displayed no SIB during the first novel probe, his rate of SIB during the second probe exceeded the treatment criterion (0.73

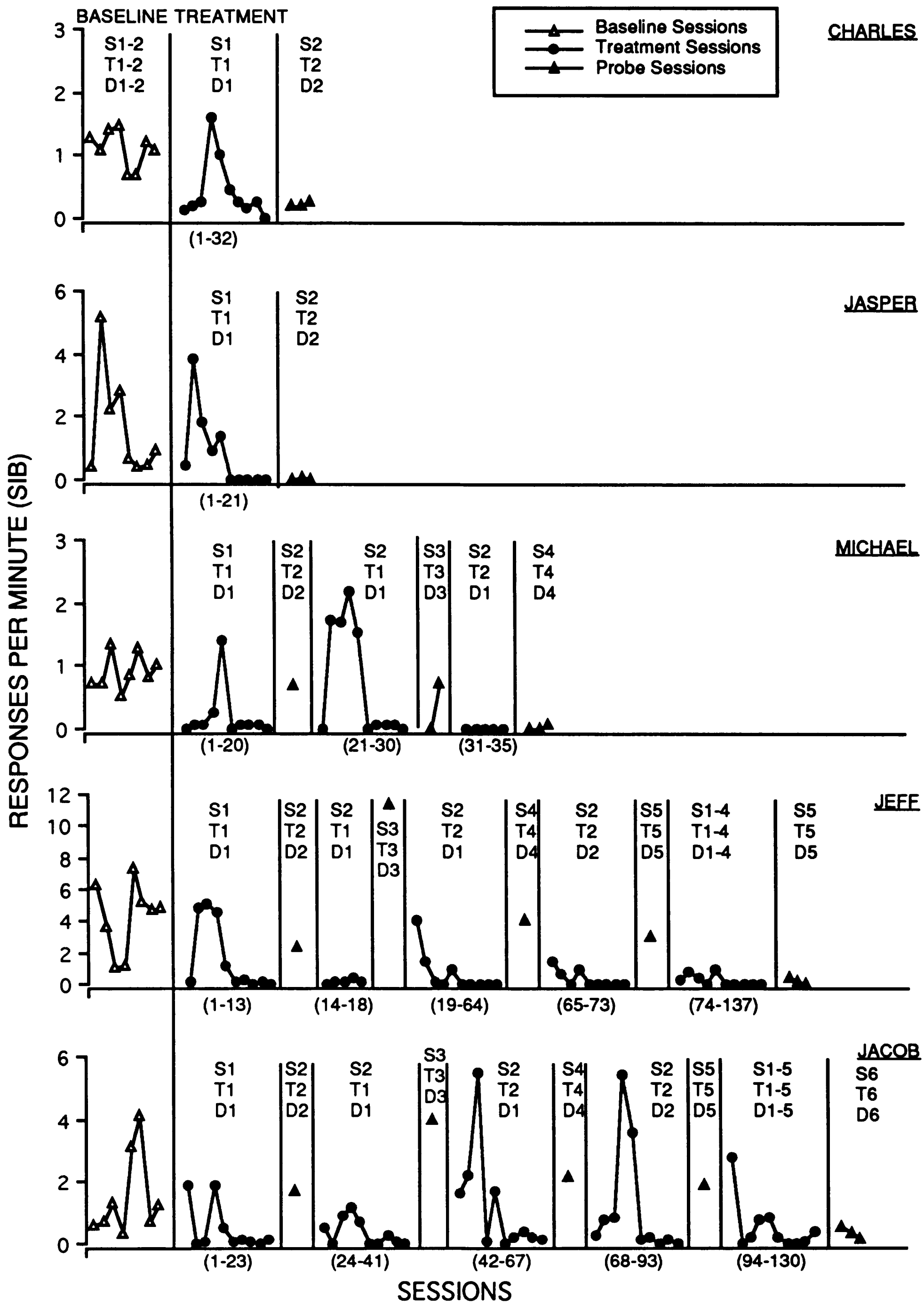

Figure 3. Responses per minute of SIB during baseline, treatment, and probe sessions.

responses per minute). Treatment was again resumed, with the setting remaining the same as in the second treatment condition (S2) and the demand set remaining the same as the original treatment condition (D1), but with a new therapist (T2). During this treatment condition (S2T2D1), criterion was reached immediately in the first five sessions (no occurrences of SIB). Novel stimulus parameters (S4T4D4) were probed following this treatment, and the generalization criterion was met with three consecutive probes at near-zero levels of SIB. Thus, Michael showed generalization following treatment across two stimulus parameters (setting and therapist).

Jeff's baseline showed variable rates of SIB across the eight sets of stimulus parameters (M = 4.27 responses per minute; range, 1 to 7.3). He did not, however, show any consistent differences in rate of SIB across the three stimulus parameters (e.g., high rates with a specific set of demands). The original treatment (S1T1D1) took 13 sessions to complete (mean rate of SIB = 2.0 responses per minute; range, 0 to 5.1). The subsequent novel probe (S2T2D2) showed an increase in rate of SIB (2.33 responses per minute). When treatment was resumed in a new setting, with the therapist and demand set the same as before (S2T1D1), criterion was reached in the first five sessions (M = 0.13 responses per minute; range, 0 to 0.33). This could have been an indication that one (or both) of the other two novel parameters (T2 or D2) was responsible for the lack of generalization seen in the novel probe (S2T2D2). The next novel probe (S3T3D3) showed even higher rates of SIB (11.3 responses per minute) and subsequent training with a novel therapist, in the second treatment setting, and with the original demand set (S2T2D1) took 44 sessions to meet criterion (mean SIB during the last 10 sessions = 0.65 responses per minute; range, 0 to 4.0). Another probe (S4T4D4) was conducted and again showed that generalization to novel parameters had not occurred (SIB = 4.0 responses per minute). The next condition, in which treatment was implemented with a new demand set but with the same setting and therapist as in the previous condition (S2T2D2), took eight sessions to complete (mean rate of SIB = 0.31 responses per minute; range, 0 to 1.47). Novel stimulus parameters (S5T5D5) were probed again and showed a high rate of SIB (3.0 responses per minute). For generalization to finally occur, Jeff required training for novelty by varying the stimulus parameters (S1 through S4, T1 through T4, D1 through D4). This was accomplished by randomly varying different combinations of settings, therapists, and demand sets (e.g., S1T2D3, S2T4D1, S4T3D2, etc.) during each session until the treatment criterion of five consecutive sessions with SIB below 0.5 responses per minute occurred. This treatment required 63 sessions to complete (mean SIB during the last 10 sessions = 0.23 responses per minute; range, 0 to 0.93). The last set of probes (S5T5D5) showed two sessions at 0 responses per minute and one session at 0.33 responses per minute (a sixth set of demands was not obtained for Jeff prior to the study; therefore, the fifth set was probed again).

Jacob's data showed results similar to those obtained for Jeff. During baseline, Jacob's rate of SIB was variable across the eight sets of stimulus parameters (M = 1.27 responses per minute; range, 0.6 to 4.1). He did not, however, show any consistent differences in rate of SIB across the three stimulus parameters. The original treatment (S1T1D1) took 23 sessions to complete (mean SIB during the last 10 sessions = 0.5 responses per minute; range, 0 to 1.87). His subsequent novel probe (S2T2D2) showed a high rate of SIB (1.7 responses per minute). Treatment was then implemented in a new setting, with the therapist and demand set the same as in the original treatment (S2T1D1). This treatment took 17 sessions to complete (mean SIB during the last 10 sessions = 0.37 responses per minute; range, 0 to 1.2). When a second novel probe was conducted (S3T3D3), a high rate of SIB was observed (4.0 responses per minute). Treatment was then resumed in the second setting, with a new therapist and the original demand set (S2T2D1), and took 26 sessions to complete (mean SIB during the last 10 sessions = 1.2 responses per minute; range, 0 to 5.5). When the third novel probe (S4T4D4) was conducted, SIB again returned to a high rate (2.13 responses per minute). Therefore, treatment was implemented in the same setting and with the same therapist as in the previous condition, but with a new demand set (S2T2D2). This condition took 23 sessions to complete (mean SIB during the last 10 sessions = 0.53 responses per minute; range, 0 to 2.8). When the fourth novel probe was conducted (S5T5D5), a high rate of SIB was observed (1.9 responses per minute). The training for novelty condition (varying all stimulus parameters, S1 through S5, T1 through T5, D1 through D5) was then implemented and required 37 sessions to complete (mean SIB during the last 10 sessions = 0.5 responses per minute; range, 0 to 2.8). Generalization finally occurred during Jacob's fifth novel probe (S6T6D6).

Compliance

Figure 4 shows the percentage compliance for each subject across treatment and probe conditions. Data are shown for the last 10 sessions of each treatment condition, and the numbers of sessions required to meet criterion in each are indicated in parentheses under each subject's graph.

Charles showed very little compliance during baseline, with five of the eight sessions at 0% compliance (*M* = 1.3%). During treatment (S1T1D1), he showed a slight increase in compliance (mean compliance for the last 10 sessions = 15.6%). There was a decrease in compliance (*M* = 6.7%) during his only novel probe (S2T2D2).

Jasper's compliance was variable during baseline (*M* = 41.5%; range, 6.5% to 73.3%), increased noticeably during the S1T1D1 treatment (mean for the last 10 sessions = 86%; range, 43.3% to 100%), but decreased again (*M* = 47.1%) during the S2T2D2 probes.

Michael's data showed negligible improvement in compliance over the course of the study. His compliance was zero throughout baseline and increased to a mean of only 7.3% during the last 10 sessions of S1T1D1 treatment. Compliance decreased thereafter and remained low until the final S4T4D4 probe, during which it averaged 8.9%.

Jeff's compliance during baseline ranged from 10% to 33.3% (*M* = 22.3%) and increased during S1T1D1 treatment (mean compliance for the last 10 sessions = 64.1%; range, 23.3% to 93.3%). When a novel S2T2D2 probe was conducted, compliance decreased to 26.7%. This pattern was repeated during subsequent conditions, increasing during treatment (S2T1D1, S2T2D1, S2T2D2, S1 through S4, T1 through T4, D1 through D4), and decreasing during novel probes (S3T3D3, S4T4D4, S5T5D5). When SIB treatment effects at last showed generalization during the final probes, Jeff's compliance increased to a mean of 75.2%.

Jacob's compliance during baseline was low (*M* = 4.6%; range, 0% to 13.3%). During the last 10 sessions of treatment in S1T1D1, his mean compliance increased to 26.3% (range, 13.3% to 46.7%). When the first novel probe (S2T2D2) was conducted, however, Jacob's compliance decreased to zero. Little or no change in compliance was observed during the next two treatment (S2T1D1, S2T2D1) and probe (S3T3D3, S4T4D4) conditions. Jacob's compliance gradually increased when treatment was implemented in S2T2D2 and was maintained during the subsequent probe (S5T5D5). During the last 10 sessions of the training for novelty condition (S1 through S5, T1 through T5, D1 through D5), Jacob's mean compliance was 29% (range, 0% to 83.3%). The final novel probe (S6T6D6) showed a mean percentage compliance of 27.8% (range, 16.7% to 36.7%).

DISCUSSION

This study illustrates a direct method for assessing and programming generalized reduction in escape-maintained SIB across multiple stimulus parameters (settings, therapists, and task demands) subsequent to treatment. Results showed highly idiosyncratic differences in generalization across subjects. Charles and Jasper showed complete generalization during the first novel probe and required no additional treatment. Michael required additional treatment with two new stimulus parameters (setting and therapist) before generalization occurred. Finally, Jeff and Jacob showed a complete lack of generalization even when treatment was extended across all three stimulus parameters (S2T2D2). Both individuals required treatment under continually varying stimulus conditions for a substantial number of sessions before generalization finally was observed.

Examination of the numbers of sessions required to meet the treatment criterion also revealed individual differences during treatment. For example, Michael met criterion immediately (in the first five

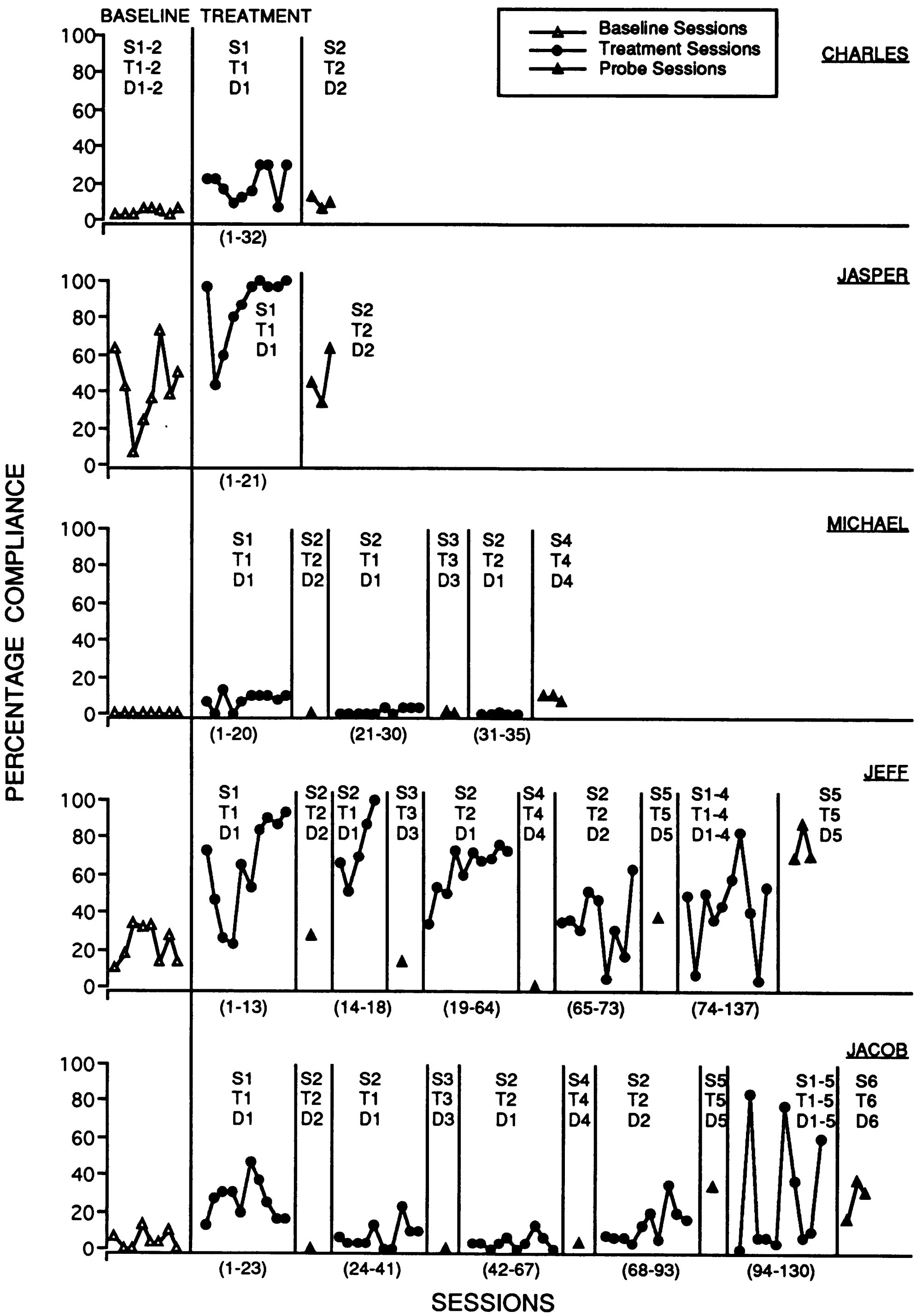

Figure 4. Percentage compliance during baseline, treatment, and probe sessions.

sessions) when treatment was implemented with a new therapist (S2T2D1), and his SIB remained low during the subsequent probe (S4T4D4). This finding suggested that the stimulus parameter affecting generalization was either the therapist or setting (both of which received treatment), but not the demand set (which remained untreated). Similarly, Jeff met criterion immediately when treatment was implemented in a new setting, with the therapist and demand set remaining the same as in the original treatment (S2T1D1). Therefore, setting may not have been the relevant stimulus parameter that prevented generalization in the immediately preceding novel probe (S2T2D2).

The compliance data for Jasper and Jeff showed increases during the treatment phases that were roughly correlated with decreases in SIB; these data complement findings of previous research examining response covariation among compliant and inappropriate behaviors (e.g., Parrish, Cataldo, Kolko, Neef, & Egel, 1986). There were, however, decreases in compliance when novel stimulus parameters were probed. This is not surprising, because one would not expect immediate increases in compliance when novel tasks are presented. Jeff's highest percentage of compliance was observed during his final novel probes, which perhaps could be attributed to his extensive exposure to similar types of tasks during the previous treatment conditions. Jacob's compliance also showed an increase toward the end of the study, although his data were less consistent than those for Jasper and Jeff. Finally, generalized reductions in SIB for Charles and Michael were associated with little or no improvement in compliance. These subjects also showed the least amount of compliance during the preexperimental demand assessment and during baseline, and were reported to rarely comply with any instructions in their home setting. For these 2 individuals, almost any instruction was an aversive event, and tolerance to the training context represented a major improvement in their behavior.

Although this study's approach to generalization programming seems counterintuitive to some unstructured training protocols designed to promote generalization (e.g., loose training, Stokes & Baer, 1977, or general case training, Day & Horner, 1989), the experimental analysis of stimulus control to identify parameters that may need manipulation during training seems important to a thorough study of generalization processes. Research aimed at developing methods to identify stimuli that acquire control during treatment to reduce inappropriate behavior and then to program for appropriate generalization is long overdue. This study represents one such method. Although the use of invariant stimuli during initial treatment may have increased the likelihood of those stimuli acquiring selective control, this procedure was necessary for the subsequent assessment of generalization using novel stimulus probes. In addition, control of extraneous stimuli during initial attempts to reduce inappropriate behavior may be necessary to determine the effectiveness of new interventions. That is, failure to decrease inappropriate behavior when treatment is conducted under varying stimulus conditions may be a function of ineffective treatment or a lack of generalization.

Each of the three stimulus parameters examined in this study was actually comprised of multiple components of other stimuli, and determination of the specific components that affected generalization requires additional analysis. The purpose of this study was not to produce a complete methodology for examining stimulus control or generalization. Instead, our goal was to demonstrate an efficient method for assessing generalization by probing across all novel parameters, and, if necessary, by programming generalization across stimulus parameters by systematically adding new ones. This seemed to be a practical first step in addressing the difficult problem of generalized behavior reduction. Given the positive findings reported here, further analyses may be undertaken to determine which (of three or even more) variables facilitate or impede generalization. Methods to conduct such an analysis could include (a) conducting novel probes consisting of only one new parameter or the most recently treated parameter and then testing further for novelty (across other parameters) if generalization occurred; or (b) conducting a completely novel probe, as was done in this study, and, if generalization did not occur, probing each parameter to see which one impeded generalization. These methods, however, incur additional limitations. First, both procedures involve introducing novel elements into successive probes, which itself may amount to additional generalization "programming." Second, the protracted probing required by these methods could eventually compromise the "control" feature of the probes.

A number of variations could be explored based on this approach to generalization programming following treatment to reduce inappropriate behavior. Research that examines the differences in generalization that may occur when the stimulus parameters are treated in different sequences (e.g., therapist or demand set treated first rather than setting) represents one such variation. The effects of the sequence of treatment cannot be determined from the present study and require further analysis. Treating one stimulus parameter at a time and then probing novel stimuli to determine if generalization has occurred, however, seems to be a parsimonious approach.

Other stimulus conditions that may affect generalization should also be examined. Differences in generalization when demand sets are similar or dissimilar or when they contain overlapping elements, for example, could lead to new techniques for facilitating generalization. Differential effects on generalization that occur when probes are conducted in settings or with therapists having a significant history with respect to the behavior problem should also be explored. In this study, the subjects had no previous history of reinforcement (for inappropriate behavior) in the novel settings or with the novel therapists, which may have affected generalization. Further research should examine generalization in the natural settings and with the staff or parents who work or live with the individual.

In summary, there are numerous possibilities for future research on the assessment of stimulus control and the development of generalized behavioral reduction. Stimuli that acquire control during treatment of behavior disorders maintained by factors other than escape, different sequences of probes, different combinations of stimuli, and alternative methods for promoting generalization subsequent to treatment are just a few examples of relatively unexplored areas. Research on the identification of stimuli that gain control during treatment to reduce inappropriate behavior and on the elucidation of methods to promote generalization is critically important to the further development of a behavior-analytic technology that is capable of producing durable therapeutic outcomes.

REFERENCES

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

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20,* 313–327.

Corte, H. E., Wolf, M. M., & Locke, B. L. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis, 4,* 209–214.

Davis, C. A., Brady, M. P., Williams, R. E., & Hamilton, R. (1992). Effects of high-probability requests on the acquisition and generalization of responses to requests in young children with behavior disorders. *Journal of Applied Behavior Analysis, 25,* 905–916.

Day, H. M., & Horner, R. H. (1989). Building response classes: A comparison of two procedures for teaching generalized pouring to learners with severe disabilities. *Journal of Applied Behavior Analysis, 22,* 223–229.

Doyle, P. J., Goldstein, H., Bourgeois, M. S., & Nakles, K. O. (1989). Facilitating generalized requesting behavior in Brocaaas aphasia: An experimental analysis of a generalization training procedure. *Journal of Applied Behavior Analysis, 22,* 157–170.

Garcia, E., Baer, D. M., & Firestone, I. (1971). The development of generalized imitation within topographically determined boundaries. *Journal of Applied Behavior Analysis, 4,* 101–112.

Halle, J. W., & Holt, B. (1991). Assessing stimulus control in natural settings: An analysis of stimuli that acquire control during training. *Journal of Applied Behavior Analysis, 24,* 579–590.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities, 3,* 1–20.

Iwata, B. A., Pace, G. M., Kalsher, M. J., Cowdery, G. E., & Cataldo, M. F. (1990). Experimental analysis and extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis, 23,* 11–28.

Kirby, K. C., & Bickel, W. K. (1988). Toward an explicit analysis of generalization: A stimulus control interpretation. *The Behavior Analyst, 11,* 115–129.

Parrish, J. M., Cataldo, M. F., Kolko, D. J., Neef, N. A., & Egel, A. L. (1986). Experimental analysis of response covariation among compliant and inappropriate behaviors. *Journal of Applied Behavior Analysis, 19,* 241–254.

Rincover, A., & Koegel, R. L. (1975). Setting generality and stimulus control in autistic children. *Journal of Applied Behavior Analysis, 8,* 235–246.

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

Stokes, T. F., Baer, D. M., & Jackson, R. L. (1974). Programming the generalization of a greeting response in four retarded children. *Journal of Applied Behavior Analysis,* 7, 599–610.

Received June 24, 1993
Initial editorial decision August 24, 1993
Revisions received October 19, 1993; February 14, 1994
Final acceptance February 14, 1994
Action Editor, David P. Wacker