JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2011, **44,** 95–108 NUMBER 1 (SPRING 2011)

FURTHER EVALUATION OF RESPONSE INTERRUPTION AND REDIRECTION AS TREATMENT FOR STEREOTYPY

ERIN N. AHRENS AND DOROTHEA C. LERMAN

UNIVERSITY OF HOUSTON–CLEAR LAKE

TIFFANY KODAK

MUNROE-MEYER INSTITUTE

UNIVERSITY OF NEBRASKA MEDICAL CENTER

AND

APRIL S. WORSDELL AND COURTNEY KEEGAN

MAY INSTITUTE

The effects of 2 forms of response interruption and redirection (RIRD)—motor RIRD and vocal RIRD—were examined with 4 boys with autism to evaluate further the effects of this intervention and its potential underlying mechanisms. In Experiment 1, the effects of motor RIRD and vocal RIRD on vocal stereotypy and appropriate vocalizations were compared for 2 participants. In Experiment 2, the effects of both RIRD procedures on both vocal and motor stereotypy and appropriate vocalizations were compared with 2 additional participants. Results suggested that RIRD was effective regardless of the procedural variation or topography of stereotypy and that vocal RIRD functioned as a punisher. This mechanism was further explored with 1 participant by manipulating the schedule of RIRD in Experiment 3. Results were consistent with the punishment interpretation.

Key words: automatic reinforcement, punishment, response interruption and redirection, stereotypy

Stereotypy is a repetitive, topographically invariant motor or vocal response that has no apparent social function. Exhibited by many individuals with autism (Bodfish, Symons, Parker, & Lewis, 2000), stereotypy may hinder the development of new skills and impede social interactions (Dunlap, Dyer, & Koegel, 1983; Wolery, Kirk, & Gast, 1985). Research indicates that stereotypy is commonly maintained by automatic reinforcement (see Rapp & Vollmer, 2005, for a review). Although noncontingent reinforcement and differential reinforcement have been shown to reduce stereotypy maintained by automatic reinforcement (e.g., Johnson, Van Laarhoven, & Repp, 2002; Piazza, Adelinis, Hanley, Goh, & Delia, 2000; Simmons, Smith, & Kliethermes, 2003), research indicates that reinforcement-based treatments for this behavior are often ineffective in the absence of other treatment components (e.g., Britton, Carr, Landaburu, & Romick, 2002; Hanley, Iwata, Thompson, & Lindberg, 2000; Lindberg, Iwata, & Kahng, 1999; Ringdahl, Vollmer, Marcus, & Roane, 1997; Vollmer, Marcus, & LeBlanc, 1994).

Providing access to stimulating objects and activities or delivering reinforcement contingent on alternative behavior establishes a concurrent schedule of reinforcement when stereotypy continues to produce reinforcing consequences. Such an arrangement may lead to treatment

Erin N. Ahrens is now at Texana Behavior Treatment and Training Center. April Worsdell is now with Coyne and Associates.

We thank Amanda Karsten, Ryan Schweck, and Sandra Rivers for their assistance with various aspects of the study.

Correspondence concerning this article should be addressed to Dorothea C. Lerman, Psychology Department, University of Houston–Clear Lake, 2700 Bay Area Blvd. Box 245, Houston, Texas 77058 (e-mail: lerman@uhcl.edu).

doi: 10.1901/jaba.2011.44-95

failure if the alternative sources of reinforcement neither compete with nor substitute for the automatic consequences of stereotypy. Possibly for this reason, blocking stereotypy (i.e., physically interrupting the response as soon as it begins to occur) has been shown to increase the effectiveness of reinforcement-based procedures (e.g., Fellner, Laroche, & Sulzer-Azaroff, 1984; Hanley et al., 2000; Lerman, Kelley, Vorndran, & Van Camp, 2003; Sprague, Holland, & Thomas, 1997; Vollmer et al., 1994). Response blocking may function as either sensory extinction (Rincover, 1978; Smith, Russo, & Le, 1999) or punishment (Lerman & Iwata, 1996). In either case, basic research indicates that the addition of this component to a reinforcement-based intervention shifts response allocation towards the alternative source of reinforcement (Critchfield, Paletz, MacAleese, & Newland, 2003; Herrnstein, 1970).

Vocal stereotypy, however, requires special consideration because the sensory consequences may not be readily blocked or modified through physical intervention. Ahearn, Clark, MacDonald, and Chung (2007) recently extended a variation of response blocking to the treatment of vocal stereotypy. The procedure, called *response interruption and redirection* (RIRD), involved a teacher stating the child's name and asking a series of questions that required the child to exhibit a vocal response (e.g., ''How old are you?'' ''What is your mother's name?'') contingent on instances of vocal stereotypy. The goal of the procedure was to interrupt stereotypy and redirect responding to appropriate vocalizations. Results showed that RIRD was effective in reducing the level of vocal stereotypy and increasing appropriate vocalizations for four children with autism. The authors suggested that the interruption component of the procedure reduced the reinforcing consequences of vocal stereotypy (i.e., functioned as sensory extinction) or RIRD functioned as punishment. The authors also suggested that appropriate vocalizations increased because the therapist delivered praise and honored requests (when possible) contingent on this behavior.

Results of Ahearn et al. (2007) are noteworthy because few studies have examined treatments for vocal stereotypy. Nonetheless, RIRD may not be a viable treatment option for individuals with a limited vocal repertoire or for those who are generally noncompliant with requests or questions. For such individuals, questions that require vocal responses would be unlikely to interrupt and redirect responding. As an alternative, delivering contingent demands that require motor responses (e.g., clapping hands, touching facial features) may be more effective for at least two reasons. First, caregivers can use progressive prompting hierarchies (e.g., model prompts followed by physical guidance) to ensure compliance with requests for motor responses, but they cannot do so with requests for vocal responses. Second, physical prompts may function as punishment for some individuals (Iwata, Pace, Cowdery, & Miltenberger, 1994). In fact, the literature on overcorrection has demonstrated the efficacy of contingent demands for reducing stereotypy (e.g., Foxx & Azrin, 1973; Wells, Forehand, Hickey, & Green, 1977). Moreover, results of a recent study showed that requiring motor responses during RIRD was highly effective for reducing vocal stereotypy (Cassella, Sidener, Sidener, & Progar, 2011).

RIRD may be most effective in reducing stereotypy if the form of the directed response matches that of stereotypy (Ollendick, Matson, & Martin, 1978). That is, RIRD that requires vocal responses (hereafter called vocal RIRD) may be more effective than RIRD that requires motor responses (hereafter called motor RIRD) in interrupting vocal stereotypy and blocking or reducing the sensory consequences of the behavior. In a similar manner, motor RIRD may be more effective than vocal RIRD in interrupting motor stereotypy and blocking or reducing the sensory consequences of the

behavior. However, if RIRD functions as punishment, the form of RIRD would be irrelevant (e.g., Epstein, Doke, Sajwaj, Sorrell, & Rimmer, 1974; Sharenow, Fuqua, & Miltenberger, 1989).

Given the limited research on treatments for vocal stereotypy in general and on RIRD for vocal stereotypy in particular, it seems beneficial to evaluate further the effectiveness of this procedure. We were also interested in identifying the mechanisms that underlie the reductive effects of RIRD on responding. In the first experiment, we compared the effects of vocal RIRD to those of a procedure that resembled vocal RIRD but did not require vocal responses (i.e., contingent demands that required motor movements). If the forms of RIRD were similarly effective for vocal stereotypy, the results would broaden the generality of the procedure and suggest that it functions as punishment. In the second experiment, we conducted the comparison with two participants who engaged in both motor and vocal stereotypy. The goal was to extend the findings of Experiment 1 by further evaluating the potential punishment function of RIRD and to identify possible advantages of matching the form of the directed response to the form of stereotypy. Finally, in the third experiment, we used the procedures described by Lerman and Iwata (1996) to systematically evaluate the mechanisms that underlie the effectiveness of RIRD.

GENERAL METHOD

Participants and Settings

Four children who had been diagnosed with autism participated. The children were eligible for this study because they had at least some functional speech and they engaged in motor or vocal stereotypy that interfered with academic programs. In addition, functional analyses conducted prior to the study using the procedures described by Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994) suggested that the targeted topographies of stereotypy were maintained by automatic reinforcement (data available from the first author). None of the participants received medication at the time of the study or had any known sensory or physical deficits. Bobby was a 6-year-old boy who was receiving home-based applied behavior-analytic services. He communicated with three- to five-word phrases mainly to request items and could answer a number of wh–questions (e.g., ''What is your name?'' What is your address?''). He described objects, people, or places in sight using up to three words when asked. He returned but did not initiate social greetings, and he imitated most vocalizations. His vocal stereotypy consisted of repeating movie lines, words, and nonspeech sounds. He participated in Experiment 1.

Hal was a 4-year-old boy who was receiving early intervention services at a hospital-based clinic. His communication consisted primarily of one-word mands (e.g., ''music'') and tacts (e.g., ''rug''). He imitated most vocalizations when prompted (e.g., ''say cookie'') but did not answer social questions or complete fill-in-the-blank statements. His vocal stereotypy consisted of a string of sounds stated repeatedly (e.g., ''oh-be-die-ay-do''). He participated in Experiments 1 and 3.

The participants in Experiment 2 were two boys who attended a day-treatment center that provided intensive behavior-analytic intervention. Glen was a 5-year-old boy who communicated vocally using four- to five-word phrases; answered social questions (e.g., age, parents' names, sibling's name, name of school); and spontaneously labeled items, activities, and people. His vocal stereotypy consisted of reciting movie lines and repeating words and phrases unrelated to the situation. His motor stereotypy consisted of hand waving, stomping his feet, and jumping up and down. He received auditory integration training for 10 days starting at the beginning of the reversal to baseline.

David was a 4-year-old boy. His communication consisted primarily of one to two words to request items. He answered some social questions (e.g., his age, name, where he went to school), responded to some fill-in-the-blank statements (e.g., "a cow says —," "you see with your —," "a bird flies in the —"), and imitated most vocalizations ("say 'cat,'" "say 'ball'"). His vocal stereotypy consisted of repeated words and noises. His motor stereotypy consisted of arm flapping, covering his ears, and body rocking.

Sessions for Bobby took place at home in his bedroom, which was approximately 3.6 m by 2.4 m and contained two twin beds, a chair, and a table with no materials on it. Sessions for Hal took place in a private therapy room that contained a table, chairs, and an opaque plastic tub filled with preferred items. Sessions for Glen and David were conducted in a day-treatment center in a room (4 m by 3 m) that contained a video camera, three to four leisure items, and a table with two chairs.

Response Measurement and Interobserver Agreement

Vocal stereotypy was defined as any nonfunctional or noncontextual speech vocalization that was not scored as an appropriate vocalization. Topographies of vocal stereotypy included repetitive sounds and phrases, singing, and recited movie lines. *Appropriate vocalizations* were defined as any vocalizations that were contextually appropriate and not prompted by the therapist, including mands (stating the name of an item that was out of sight) and tacts (stating the name of an item within 0.9 m of the participant). Requested vocal responses that occurred during RIRD sequences were not scored as appropriate vocalizations (see compliance below). Motor stereotypy (Experiment 2 only) included *hand flapping,* defined as rapid movement of the hand in a back and forth motion; *body rocking,* defined as the forward and backward movement of the upper body; and *clapping,* defined as the rapid movement of the hands hitting together. *Compliance* during RIRD (Glen, David, and Hal) was defined as exhibiting the requested vocal or motor response following the initial instruction or the first prompt. Data on vocal stereotypy were collected using 10-s partial-interval recording (Bobby, Glen, and David) or duration recording (Hal). These data were converted to percentage of intervals (partial-interval recording) or percentage of session time (duration recording) after dividing the total number of intervals scored (including intervals during which the therapist delivered the RIRD instructional sequences; Experiment 2 only) or the total number of seconds by the total number of intervals or the total duration of the session. Stereotypy that occurred during the RIRD instructional sequences was not included in the data analysis for Bobby and Hal (Experiments 1 and 3). All stereotypy was included in the data analysis for Glen and David (Experiment 2) to examine the potential benefits of matching the form of RIRD to the topography of stereotypy. Data on appropriate vocalizations were scored using 10-s partial-interval recording (Bobby, Glen, and David) or frequency recording (Hal). Data on motor stereotypy were collected using 10-s partial-interval recording (Glen and David). Data on compliance (Glen, David, and Hal) were scored using frequency recording and converted to percentage of instructions during RIRD. Trained observers used laptop computers (Hal only) or paper and pencil to collect data, along with a stopwatch or MotivAider to help keep track of the intervals. Sessions for Glen and David were videotaped for data-collection purposes.

A second observer collected data independently during at least 20% (Experiments 1 and 3) and 25% (Experiment 2) of the sessions in each condition for reliability purposes. Interobserver agreement was calculated by dividing the number of 10-s intervals with agreement by the total number of intervals with agreements plus disagreements and converting the ratio to a

percentage. Agreement for stereotypy, appropriate vocalizations, and compliance (with the exception of Hal's vocal stereotypy) was defined as both observers scoring the same number of responses in the interval. Agreement for Hal's vocal stereotypy was defined as both observers scoring the onset or offset of behavior at the same second in each interval.

For Bobby, mean interobserver agreement was 92% (range, 80% to 100%) for vocal stereotypy and 96% (range, 83% to 100%) for appropriate vocalizations. For Hal, mean agreement was 82% (range, 62% to 100%) for vocal stereotypy, 98% (range, 98% to 100%) for appropriate vocalizations, and 97% (range, 94% to 100%) for compliance. For Glen, mean agreement was 87% (range, 65% to 100%) for vocal stereotypy, 95% (range, 86% to 100%) for motor stereotypy, 95% (range, 87% to 100%) for appropriate vocalizations, and 100% for compliance. For David, mean agreement was 81% (range, 64% to 100%) for vocal stereotypy, 90% (range, 67% to 100%) for motor stereotypy, 92% (range, 85% to 100%) for appropriate vocalizations, and 100% for compliance.

EXPERIMENT 1

Design and Procedure

A combined reversal and multielement design was used to compare the effects of two RIRD techniques on vocal stereotypy. A different therapist was associated with each RIRD condition for Hal. The procedures varied slightly for Bobby and Hal, because the study was conducted at different research sites.

Baseline. The therapist and participant were in the room. No programmed consequences were arranged for vocal stereotypy. If the participant engaged in appropriate vocalizations, the therapist delivered praise and any requested items or activities if possible (e.g., tickles). If the participant requested something that was not available (e.g., drink), the therapist delivered praise and told the participant that he might receive the item later. Sessions lasted 5 min. The two therapists associated with the different RIRD procedures conducted baseline sessions in an alternating fashion. Bobby participated in one to eight baseline sessions per day, usually 1 day per week. Hal participated in two to six baseline sessions per day, 2 to 4 days per week.

RIRD. Contingent on vocal stereotypy, the therapist stated the child's name in a neutral tone of voice, immediately followed by social questions or instructions that required a vocal response for vocal RIRD (e.g., "Where do you live?" "say 'Mom'") or by instructions that required a motor response for motor RIRD (e.g., "stand up"). The questions and instructions delivered during the RIRD sequences were associated with responses that the participant had mastered during his regular therapy sessions prior to the study. The therapists randomly selected from among the questions, fill-in-the-blank statements, and instructions (e.g., "What is your name?" "You sleep in a —," "clap hands," "do this"). Criterion for mastery was at least 80% independent responding on 3 of consecutive days (Bobby) or at least 90% independent responding on 2 consecutive days (Hal). During vocal RIRD, the therapist asked the question or delivered the instruction (e.g., "What color?") and waited 5 s for the participant to respond. If he did not respond within 5 s, Bobby's therapist reissued the question (e.g., "What color?"), whereas Hal's therapist provided a model prompt (e.g., "say 'ball'"). If the participant did not respond to the prompt (or second question) within 5 s, the therapist provided a prompt (e.g., "What color? green"). If he still did not respond to the prompt, the therapist initiated the next trial. If the participant responded independently or to a prompt, the therapist provided praise (e.g., "that's right," "nice job") and continued to present vocal demands until he experienced three consecutive RIRD trials without engaging in stereotypy. However, the participant was not

required to comply with the vocal demands on each trial to terminate the RIRD sequence, a procedure that deviated from that of Ahearn et al. (2007). This requirement was not included in the current study (a) to reduce the time associated with implementing RIRD, thereby increasing its practicality, and (b) to determine if the procedure would be effective even if the individual did not comply with the requests.

During motor RIRD, the therapist delivered an instruction and waited 5 s for the participant to respond. If he did not respond within 5 s, the therapist modeled the correct motor response (or delivered a physical prompt for motor imitation targets). If the participant did not respond within 5 s of the model prompt, the therapist physically guided him to engage in the correct response. If the participant responded independently or to the model prompt, the therapist delivered praise (''that's right'' or ''nice job''). The therapist continued to present the demands until he had engaged in three consecutive responses (prompted or unprompted) without engaging in stereotypy.

During both vocal and motor RIRD, the therapist responded to any appropriate vocalizations such as labeling items in the room (e.g., ''I see the camera'') or requesting an item (e.g., ''I want the book'') or the therapist's attention (e.g., ''look at me,'' ''hug me'') by providing attention or the item requested. If the item requested was not available (e.g., going out to the playground), the therapist acknowledged the request and stated that it would be available later (e.g., ''That's nice asking but maybe later''). Sessions continued for 5 min, not including time consumed by the RIRD procedure, or until 30 min had passed. None of Bobby's sessions met the early termination criterion. For Hal, the therapist terminated three sessions of vocal RIRD and two sessions of motor RIRD prior to reaching 5 min due to the 30-min limit. Bobby's therapist implemented two sessions each day (one for each treatment condition, counterbalanced across days), 1 day per week. Hal's therapist conducted two to six sessions per day, equally divided across the two conditions, 2 to 4 days per week. At least 15 min elapsed between sessions.

Results and Discussion

Results for vocal stereotypy (top) and appropriate vocalizations (bottom) are shown in Figure 1 for Bobby (left) and Hal (right). Bobby engaged in high levels of vocal stereotypy (M = 59%) and low levels of appropriate vocalizations (M = 3.3%) in baseline. The RIRD conditions produced similar reductions in vocal stereotypy (M = 31% for vocal RIRD and M = 35% for motor RIRD). Both conditions also were associated with increases in appropriate vocalizations (M = 30% for vocal RIRD, M = 28% for motor RIRD). Similar results occurred with the return to baseline and reintroduction of the treatment comparison, although appropriate vocalizations were somewhat higher under the motor RIRD condition (M = 31%) than in the vocal RIRD condition (M = 15%) in the final treatment phase.

For Hal, vocal stereotypy occurred during a moderate proportion of the session time during baseline (M = 30%). He engaged in few instances of appropriate vocalizations. Vocal RIRD and motor RIRD produced similar reductions in vocal stereotypy, although vocal RIRD appeared to be associated with slightly faster reductions in responding (M = 9.8% for vocal RIRD, M = 23% for motor RIRD). Both conditions were associated with increases in appropriate vocal behavior (M = 11.5% for vocal RIRD, M = 8.5% for motor RIRD). However, an overall decrease in appropriate vocalizations occurred across three additional reversals to baseline and replication of the treatments within multielement and reversal designs.

These findings indicate that vocal and motor RIRD were similarly effective for reducing vocal stereotypy. Most notably, both interventions were associated with increases in appropriate vocalizations, suggesting that vocal redirection

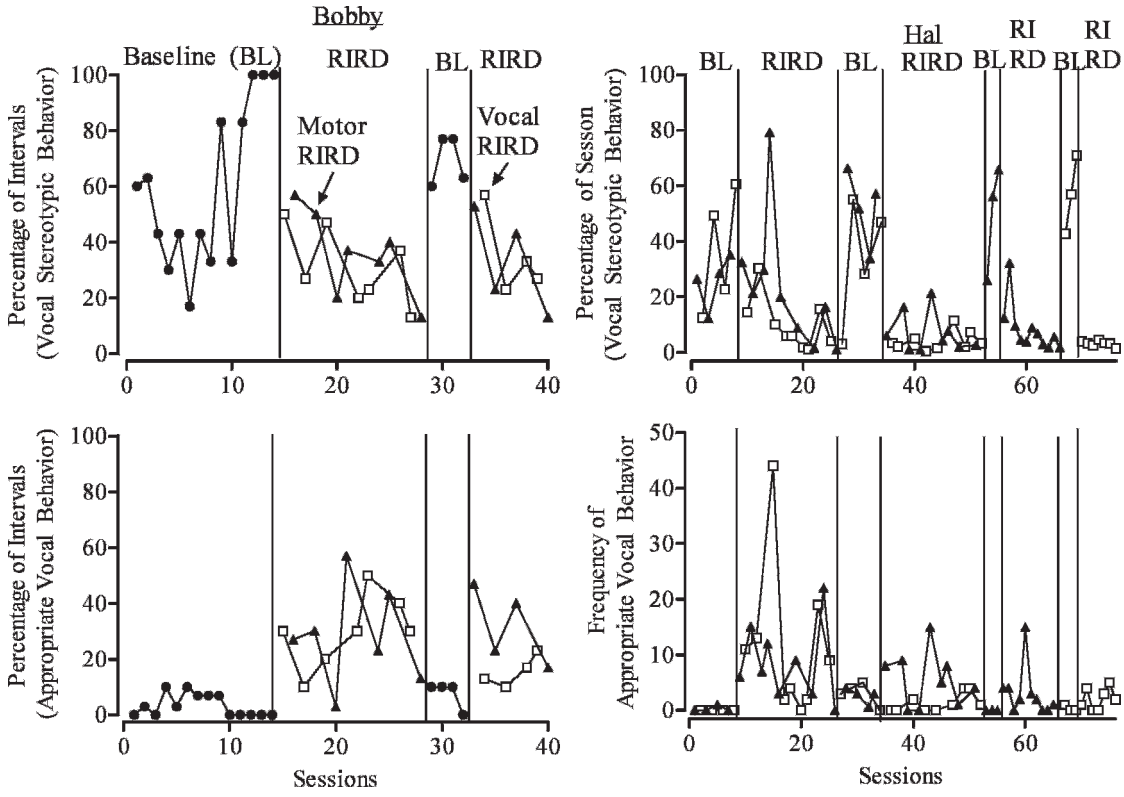

Figure 1. Percentage of intervals with vocal stereotypy (top left) and appropriate vocalizations (bottom left) for Bobby during the treatment comparison. Percentage of session time with vocal stereotypy (top right) and frequency of appropriate vocalizations (bottom right) for Hal during the treatment comparison.

per se was not necessary to produce this outcome. Covariation between inappropriate and appropriate behavior has been observed with other types of responses and interventions (e.g., Iwata, Pace, Kalsher, Cowdery, & Cataldo, 1990; Parrish, Cataldo, Kolko, Neef, & Egel, 1986). These results suggest that motor RIRD might be a viable alternative to vocal RIRD for individuals who have a limited vocal repertoire or who are often noncompliant with vocal requests. Furthermore, results appear to indicate that RIRD can be effective even if the individual is not required to comply with the requests to terminate the instructional sequence. Data on compliance to the instructions delivered during the RIRD sequences were collected for Hal only. Levels of compliance were variable but generally low, and they appeared to have no relation to the effectiveness of the interventions ($M = 15.6\%$ under vocal RIRD and $M = 19\%$ under motor RIRD).

Together, these findings suggest that vocal RIRD functioned as punishment for these participants because the form of RIRD was irrelevant to the treatment effects. If RIRD reduced stereotypy by blocking or reducing the sensory consequences of the behavior, vocal RIRD probably would be more effective than motor RIRD in treating vocal stereotypy. Given the similar outcomes for vocal and motor RIRD on vocal stereotypy, results also suggest that the form of the instructed response does not need to match the form of stereotypy. However, a more thorough evaluation of both questions would require participants who engage in both vocal and motor stereotypy. Thus, we conducted a

second experiment with two additional children who engaged in both motor and vocal stereotypy.

EXPERIMENT 2

Design and Procedure

A combined reversal and multielement design was used to the compare the effects of vocal and motor RIRD on both vocal and motor stereotypy for Glen and David. A different therapist was associated with each condition. Therapists conducted one or two sessions per day (depending on staff availability), 4 days per week.

Baseline. Procedures were identical to those in the baseline condition of Experiment 1. However, all sessions lasted 10 min. The two therapists associated with the different RIRD procedures conducted baseline sessions in an alternating fashion.

RIRD. Procedures for vocal and motor RIRD were similar to those described in Experiment 1, except that the prompts were separated by 2 s to 3 s (instead of 5 s). Furthermore, the therapist delivered RIRD sequences contingent on either vocal or motor stereotypy. The questions and instructions delivered during the RIRD sequences were associated with responses that the participants had mastered during their regular therapy sessions prior to the study. Criteria for mastery was 90% independent responding across three different therapists for 3 consecutive days. The therapist delivered instructions and questions that required a vocal response (e.g., ''What is your name?'' ''A cow says —'') during vocal RIRD and instructions that required the participant to engage in small motor movements (e.g., clapping hands, touching facial features, patting lap) during motor RIRD.

Each RIRD procedure was associated with a different therapist to reduce the possibility of interaction effects. A coin flip prior to each session determined which treatment would be implemented; however, no more than two sessions with a particular treatment were implemented consecutively. In both treatment conditions, sessions continued for 10 min, not including time consumed by the RIRD procedure. However, total session time was not permitted to exceed 30 min. Four sessions for Glen and two sessions for David met the criterion for early termination.

Results and Discussion

Results of Glen's treatment evaluation are displayed in Figure 2. He exhibited high levels of vocal stereotypy in baseline (M = 81%). When vocal RIRD was implemented, vocal stereotypy decreased slightly, with levels never falling below 40% of intervals (M = 62%). Motor RIRD resulted in a more rapid decrease in vocal stereotypy, with levels falling far below those in baseline (M = 42%). Vocal stereotypy increased to a high level during the return to baseline (M = 87%). When both RIRD procedures were reintroduced, motor RIRD again resulted in an overall greater decrease in vocal stereotypy (M = 39% for motor RIRD and M = 49% for vocal RIRD). Baseline levels of motor stereotypy were high (M = 78%). When vocal RIRD was implemented, motor stereotypy remained at baseline levels for the first few sessions and then gradually decreased (M = 45%). Motor RIRD resulted in a more rapid decrease in motor stereotypy, with levels at or below 25% of intervals during the last five sessions (M = 34%). During the return to baseline, motor stereotypy increased to levels similar to those in the original baseline (M = 76%). When RIRD was reintroduced, motor RIRD again resulted in a greater decrease in the level of motor stereotypy compared to vocal RIRD (M = 29% for motor RIRD, and M = 36% for vocal RIRD). Baseline levels of appropriate vocalizations were low (M = 21%). Levels increased during both vocal RIRD (M = 37%) and motor RIRD (M = 35%), with no differences between the two conditions. Appropriate vocalizations continued to occur at high levels during the return to baseline (M =

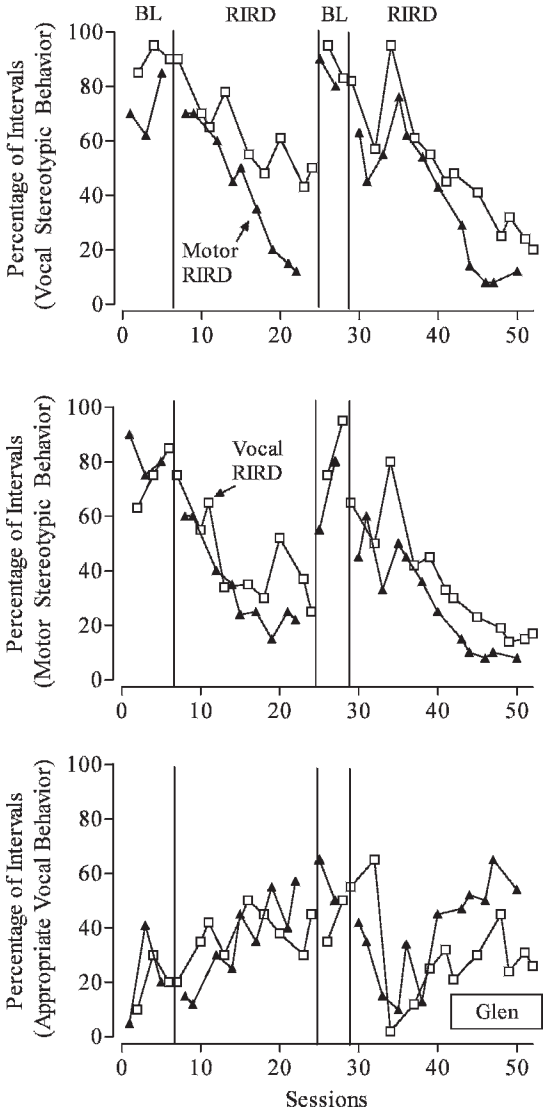

Figure 2. Percentage of intervals with vocal stereotypy (top), motor stereotypy (middle), and appropriate vocalizations (bottom) for Glen during the treatment comparison.

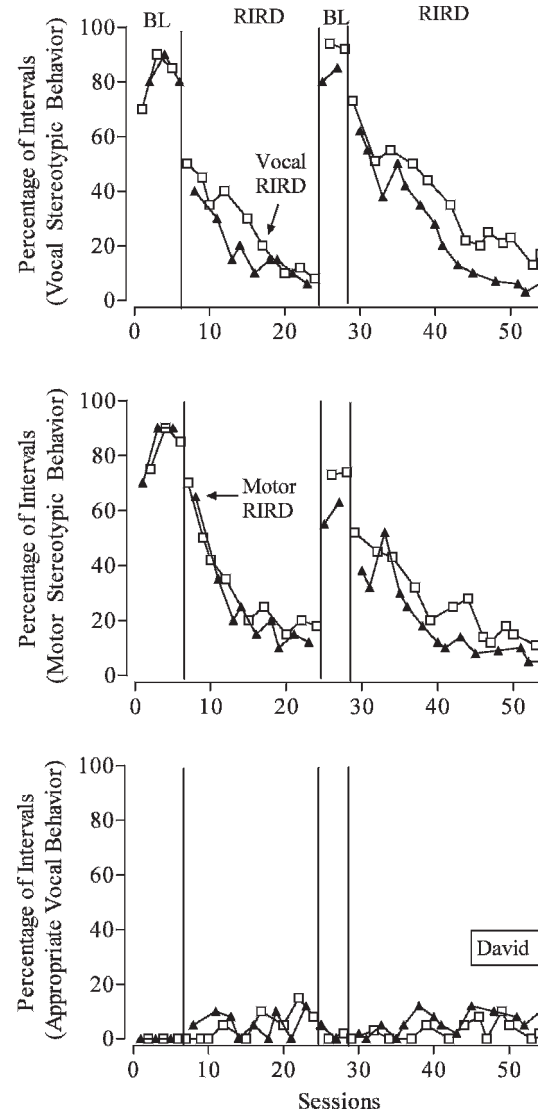

Figure 3. Percentage of intervals with vocal stereotypy (top), motor stereotypy (middle), and appropriate vocalizations (bottom) for David during the treatment comparison.

50%) and then temporarily decreased before gradually increasing when vocal RIRD was reintroduced. Levels were somewhat higher under motor RIRD (M = 38%) than under vocal RIRD (M = 31%).

Figure 3 displays the results of David's treatment evaluation. Vocal stereotypy occurred at high levels during baseline (M = 82%). Vocal stereotypy decreased to low levels under both RIRD procedures, but the reductions were slightly more rapid during motor RIRD (M = 18%) than during vocal RIRD (M = 28%). Vocal stereotypy increased to a high level during the return to baseline (M = 88%). When RIRD was reintroduced, both vocal RIRD and motor RIRD resulted in a steady decrease in vocal stereotypy, but motor RIRD (M = 27%) again produced slightly greater decreases in stereotypy

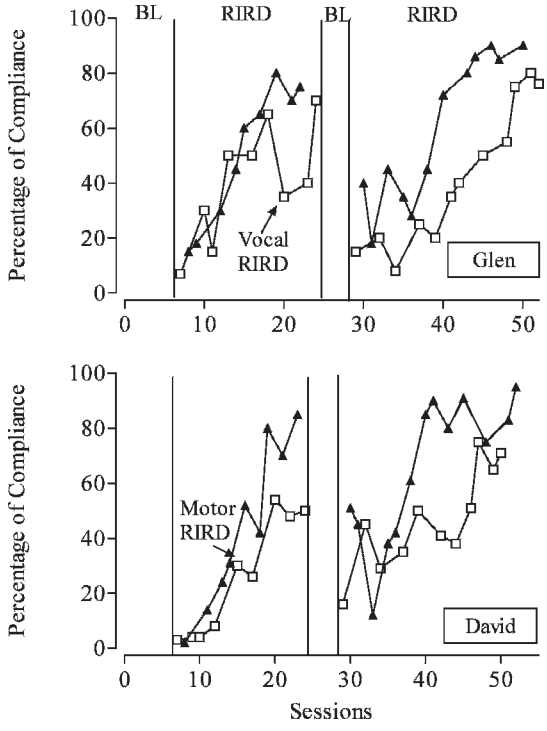

Figure 4. Percentage of compliance during the treatment comparison for Glen (top) and David (bottom).

relative to vocal RIRD (M = 34%). Baseline levels of motor stereotypy were high (M = 83%). Similar decreases in stereotypy occurred under both vocal RIRD and motor RIRD, although the overall mean levels of responding were slightly lower under the latter procedure (M = 24% for motor RIRD and M = 32% for vocal RIRD). Responding increased with the return to baseline (M = 66%), although levels were lower than those in the initial baseline phase. When both RIRD procedures were reintroduced, slightly greater reductions in motor stereotypy occurred during motor RIRD (M = 19%) than during vocal RIRD (M = 25%). No appropriate vocalizations occurred during baseline, but responding increased to levels slightly above baseline during vocal RIRD and motor RIRD (M = 5% for both procedures). Appropriate vocalizations remained low during the return to baseline (M = 2%) and then increased again when RIRD was reintroduced, with slightly higher levels during motor RIRD than during vocal RIRD (Ms = 3% for vocal RIRD and 6% for motor RIRD).

Figure 4 shows the percentage of compliance to the demands and questions delivered during motor and vocal RIRD for Glen and David. Levels of compliance increased across the treatment phases for both participants. Glen complied with a higher percentage of the demands during motor RIRD than during vocal RIRD, particularly toward the latter part of the first treatment comparison and during the second treatment comparison (overall, M = 41% for vocal RIRD and M = 55% for motor RIRD). Similar results were obtained for David (overall, M = 38.5% for vocal RIRD and M = 55% for motor RIRD).

These results showed that both variations of RIRD were effective in reducing vocal and motor stereotypy. The procedures also were associated with similar increases in appropriate vocalizations. These findings replicate those reported by Ahearn et al. (2007) and further show that the treatment can be effective even when the individual is not required to comply with the requests to terminate the RIRD sequences. In contrast to the findings obtained with the participants in Experiment 1, motor RIRD was slightly more effective than vocal RIRD in reducing both forms of stereotypy. Thus, either variation of RIRD would be effective in treating stereotypy and increasing appropriate vocalizations. Moreover, matching the form of the RIRD to the topography of stereotypy does not appear to convey additional advantages when the goal is to interrupt and reduce responding. In fact, a variation of RIRD that requires a series of small motor movements (rather than vocal responses) may produce more rapid reductions in stereotypy.

These results, which are consistent with research findings on overcorrection (e.g., Epstein et al., 1974; Foxx & Azrin, 1973), further suggest that vocal and motor RIRD functioned

as punishment for the participants in this study. If so, motor RIRD may have been more effective than vocal RIRD in reducing stereotypy because the requests for motor movements or the physical prompts delivered contingent on noncompliance functioned as more potent punishers than the requests and vocal prompts for vocal responses. The participants also may have engaged in higher levels of compliance under motor RIRD to avoid the physical prompts. The process responsible for the effects of vocal RIRD on responding was examined more directly in Experiment 3 by using the procedures described by Lerman and Iwata (1996) to distinguish between extinction and punishment.

EXPERIMENT 3

Following the comparison of vocal and motor RIRD in Experiment 1, we implemented a treatment integrity fading procedure with Hal to evaluate the possible operant mechanisms that were responsible for the effectiveness of treatment.

Design and Procedure

Procedures were similar to those described by Lerman and Iwata (1996). Three fading steps, each corresponding to a different proportion of responses that produced the RIRD sequence, were evaluated to determine whether RIRD functioned as extinction or punishment. The rationale for the fading was as follows: If RIRD functioned as punishment, stereotypy would decrease as more responses were blocked (i.e., as the schedule of punishment became richer). Results of Lerman and Iwata, who examined the function of response blocking, were consistent with this interpretation. If RIRD functioned as extinction, vocal stereotypy would persist and possibly increase as more responses were blocked (i.e., as the schedule of automatic reinforcement was thinned). Smith et al. (1999) obtained this pattern when replicating the procedures described by Lerman and Iwata.

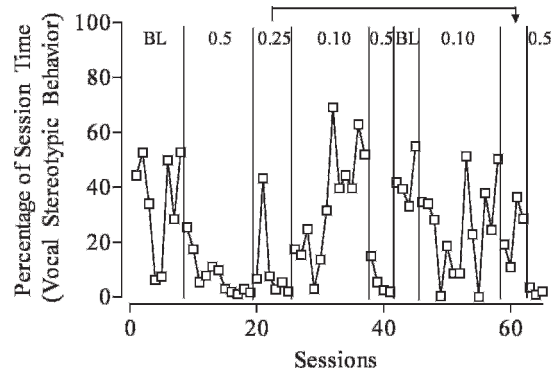

Figure 5. Percentage of session time with vocal stereotypy for Hal during the schedule manipulation in Experiment 3.

Responding during baseline and the three fading steps was evaluated in a reversal design. During baseline, the therapist provided no consequences for vocal stereotypy. Under the .5 condition, the therapist implemented the treatment following every other instance of vocal stereotypy. Thus, treatment followed 50% of the responses. Under the .25 condition, the therapist implemented the treatment following every fourth instance of vocal stereotypy (i.e., the treatment followed 25% of the responses). Under the .1 condition, treatment was implemented following the 10th occurrence of vocal stereotypy (i.e., treatment followed 10% of the responses). Thus, when transitioning across the three fading steps from .5 to .1, the schedule of automatic reinforcement would become richer while the schedule of punishment would become leaner, producing opposite effects on responding. The procedures for vocal RIRD were identical to those in Experiment 1.

Results and Discussion

The percentage of intervals with vocal stereotypy during baseline and the three fading conditions is shown in Figure 5. Moderate and variable levels of stereotypy occurred in baseline (M = 34%). Responding decreased to low levels when 50% of the responses produced RIRD (M = 8%). Stereotypy temporarily increased and then decreased when only 25%

of the responses produced RIRD (M = 11%). However, responding increased to baseline levels when only 10% of the responses were followed by RIRD (M = 34%). Reintroduction of the .5 condition produced low levels of stereotypy. Following a return to baseline, the .1, .25, and .5 conditions were replicated. Levels of stereotypy were negatively related to the percentage of responses that produced RIRD (M = 25% for .1, M = 23% for .25, and M = 2% for .5). These results are consistent with a punishment interpretation of RIRD. That is, as more responses were blocked (i.e., as the schedule of punishment increased), the treatment was more effective in reducing stereotypy. If vocal RIRD functioned as extinction, the behavior should have been maintained when 50% of the responses produced RIRD. Instead, Hal's vocal stereotypy decreased relatively quickly. Lerman and Iwata (1996) obtained a similar outcome when they examined the function of response blocking with one participant. Conversely, in Smith et al. (1999), responding was maintained under the .5 condition and gradually increased as more responses were blocked, consistent with an extinction interpretation.

GENERAL DISCUSSION

Results suggest that RIRD is highly effective for reducing stereotypy, regardless of the topography of stereotypy or requested response and regardless of whether the participant is required to engage in the vocal or motor responses during the RIRD sequences. These findings broaden the generality of the RIRD procedure for reducing vocal stereotypy, because individuals may not need to possess a vocal repertoire or be highly compliant to benefit from the intervention.

In Ahearn et al. (2007), the participants were required to comply with three consecutive requests to terminate the RIRD sequences. This criterion may render RIRD impractical for individuals who are often noncompliant to requests. Motor RIRD circumvents this problem by incorporating physical prompts to ensure compliance. Nonetheless, we also found that vocal RIRD was effective even for participants who were usually noncompliant. Thus, the less stringent criterion used in this study (i.e., refraining from stereotypy during three consecutive demand trials) is a viable modification that might increase the practicality and generality of vocal RIRD. The intervention was still somewhat labor intensive for the therapists in this study. However, by the end of the treatment evaluation (i.e., during the last three treatment sessions of each RIRD condition), Bobby's mean overall session length was 7.6 min (during vocal RIRD) and 7.4 min (during motor RIRD) with intervention time added to each 5-min treatment session. Glen's therapists spent an average of 14 min (for vocal RIRD) and 12 min (for motor RIRD) implementing 10-min treatment sessions. Results were similar for David, whose mean overall session length was 13 min (during vocal RIRD) and 12 min (during motor RIRD) with intervention time added to each 10-min session. Conversely, the therapist for Hal, who also participated in 5-min sessions, spent a mean of just 5.2 min and 5.3 min implementing the last three treatment sessions.

Together, results of the three experiments suggested that vocal RIRD functioned as punishment for the four participants in this study. If vocal RIRD functioned as extinction, it is unlikely that both forms of RIRD would have produced similar outcomes. Neither the form of RIRD nor the topography of the stereotypy seemed relevant to the efficacy of the procedure. Results were also consistent with punishment effects when the proportion of responses that produced the consequence was systematically manipulated for one participant using the procedures described by Lerman and Iwata (1996). The outcomes were not identical to those of Lerman and Iwata, who obtained

consistent decreases in responding even when just 25% and 10% of responses were blocked. However, other studies have shown that intermittent punishment may not be as effective as continuous punishment (e.g., Clark, Rowbury, Baer, & Baer, 1973). Thus, RIRD appears to be procedurally and functionally similar to the class of punishment procedures that involve the use of contingent effort (e.g., overcorrection).

Nonetheless, a number of limitations should be considered. Although both RIRD procedures were associated with increases in appropriate vocalizations, the outcomes were somewhat unclear for Glen and Hal, and David's appropriate vocalizations remained relatively low throughout the treatment. Furthermore, all participants in this study had a vocal repertoire; thus, the increase in appropriate vocalizations may not generalize to participants without this repertoire. Interaction effects could account for the reductions in vocal stereotypy and the increases in appropriate vocalizations under motor RIRD, because this procedure was rapidly alternated with vocal RIRD within a multielement design. It should be noted, however, that similar reductions were obtained when the procedures were compared in a reversal design for Hal. Despite the rapid reductions in responding, RIRD also may be difficult to implement in some applied settings. Future research is needed on the efficacy and practicality of this treatment when implemented in more naturalistic settings and throughout an individual's daily activities. For example, results of Experiment 3 indicated that Hal's treatment would be unsuccessful unless caregivers delivered the RIRD sequence following at least half of his stereotypic responses.

Interobserver agreement for vocal stereotypy also was somewhat low for some participants and sessions. This may have occurred for Glen and David because the second observer scored sessions from video and occasionally had difficulty hearing the content of the vocalizations on the video (due to the quality of the recording and the natural volume of the participants' voices). In addition, some objects in the room were obscured on the video, making it difficult to distinguish between a tact (e.g., labeling a picture in a book, which would be an appropriate vocalization) and vocal stereotypy. The lower levels of agreement for Hal's vocal stereotypy were likely related to the stringent criteria associated with measuring exact agreement (i.e., required both data collectors to score the onset or offset of behavior at the same second in an interval for the interval to be scored as an agreement). Hal's vocal stereotypy often consisted of frequent, short intervals of behavior. A review of the raw data indicated that disagreements were most commonly associated with a difference of 1 s in scoring the onset or offset of vocal stereotypy.

In conclusion, these findings suggest that RIRD, like other variations of overcorrection, may be effective in reducing stereotypy. This is particularly noteworthy for vocal stereotypy because few interventions have been developed to treat this behavior. Moreover, vocal RIRD is relatively nonintrusive and may be a viable option when reinforcement-based procedures alone are ineffective in reducing problem behavior.

REFERENCES

Ahearn, W. H., Clark, K. M., MacDonald, R. P., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis*, *40*, 263–275.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders*, *30*, 237–243.

Britton, L. N., Carr, J. E., Landaburu, H. J., & Romick, K. S. (2002). The efficacy of noncontingent reinforcement as treatment for automatically reinforced stereotypy. *Behavioral Interventions*, *17*, 93–103.

Cassella, M. D., Sidener, T. M., Sidener, D. W., & Progar, P. R. (2011). Response interruption and redirection for vocal stereotypy in children with autism: A systematic replication. *Journal of Applied Behavior Analysis*, *44*, 169–173.

Clark, H. B., Rowbury, T., Baer, A. M., & Baer, D. M. (1973). Timeout as a punishing stimulus in continuous and intermittent schedules. *Journal of Applied Behavior Analysis*, *6*, 443–455.

Critchfield, T. S., Paletz, E. M., MacAleese, K. R., & Newland, M. C. (2003). Punishment in human choice: Direct or competitive suppression? *Journal of the Experimental Analysis of Behavior*, *80*, 1–27.

Dunlap, G., Dyer, K., & Koegel, R. L. (1983). Autistic self-stimulation and intertrial interval duration. *American Journal of Mental Deficiency*, *88*, 194–202.

Epstein, L. H., Doke, L. A., Sajwaj, T. E., Sorrell, S., & Rimmer, B. (1974). Generality and side effects of overcorrection. *Journal of Applied Behavior Analysis*, *7*, 385–390.

Fellner, D. J., Laroche, M., & Sulzer-Azaroff, B. (1984). The effects of adding interruption to differential reinforcement on targeted and novel self-stimulatory behaviors. *Journal of Behavior Therapy and Experimental Psychiatry*, *15*, 315–321.

Foxx, R. M., & Azrin, N. H. (1973). The elimination of autistic self-stimulatory behavior by overcorrection. *Journal of Applied Behavior Analysis*, *6*, 1–14.

Hanley, G. P., Iwata, B. A., Thompson, R. H., & Lindberg, J. S. (2000). A component analysis of "stereotypy as reinforcement" for alternative behavior. *Journal of Applied Behavior Analysis*, *33*, 285–297.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243–266.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*, 197–209. (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2*, 3–20, 1982)

Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis*, *27*, 131–144.

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*, 11–27.

Johnson, J. W., Van Laarhoven, T., & Repp, A. C. (2002). Effects of stereotypy and other challenging behavior of matching rates of instruction to free-operant rates of responding. *Research in Developmental Disabilities*, *23*, 266–284.

Lerman, D. C., & Iwata, B. A. (1996). A methodology for distinguishing between extinction and punishment effects associated with response blocking. *Journal of Applied Behavior Analysis*, *29*, 231–234.

Lerman, D. C., Kelley, M. E., Vorndran, C. M., & Van Camp, C. M. (2003). Collateral effects of response blocking during the treatment of stereotypic behavior. *Journal of Applied Behavior Analysis*, *36*, 119–123.

Lindberg, J. S., Iwata, B. A., & Kahng, S. W. (1999). On the relation between object manipulation and stereotypic self-injurious behavior. *Journal of Applied Behavior Analysis*, *32*, 51–62.

Ollendick, T. H., Matson, J. L., & Martin, J. E. (1978). Effectiveness of hand overcorrection for topographically similar and dissimilar self-stimulatory behavior. *Journal of Experimental Child Psychology*, *25*, 396–403.

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.

Piazza, C. C., Adelinis, J. D., Hanley, G. P., Goh, H., & Delia, M. D. (2000). An evaluation of the effects of matched stimuli on behaviors maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*, *33*, 13–27.

Rapp, J. T., & Vollmer, T. V. (2005). Stereotypy: I. A review of behavioral assessment and treatment. *Research in Developmental Disabilities*, *26*, 527–547.

Rincover, A. (1978). Sensory extinction: A procedure for eliminating self-stimulatory behavior in developmentally disabled children. *Journal of Abnormal Child Psychology*, *6*, 299–310.

Ringdahl, J. E., Vollmer, T. R., Marcus, B. A., & Roane, H. S. (1997). An analogue evaluation of environmental enrichment: The role of stimulus preference. *Journal of Applied Behavior Analysis*, *30*, 203–216.

Sharenow, E. L., Fuqua, R. W., & Miltenberger, R. G. (1989). The treatment of muscle tics with dissimilar competing response practice. *Journal of Applied Behavior Analysis*, *22*, 35–42.

Simmons, J. N., Smith, R. G., & Kliethermes, L. (2003). A multiple-schedule evaluation of immediate and subsequent effects of fixed-time food presentation on automatically maintained mouthing. *Journal of Applied Behavior Analysis*, *36*, 541–544.

Smith, R. G., Russo, L., & Le, D. D. (1999). Distinguishing between extinction and punishment effects of response blocking: A replication. *Journal of Applied Behavior Analysis*, *32*, 367–370.

Sprague, J., Holland, K., & Thomas, K. (1997). The effect of noncontingent sensory reinforcement, contingent sensory reinforcement, and response interruption on stereotypical and self-injurious behavior. *Research in Developmental Disabilities*, *18*, 61–77.

Vollmer, T. R., Marcus, B. A., & LeBlanc, L. (1994). Treatment of self-injury and hand mouthing following inconclusive functional analyses. *Journal of Applied Behavior Analysis*, *27*, 331–344.

Wells, K. C., Forehand, R., Hickey, K., & Green, K. D. (1977). Effects of a procedure derived from the overcorrection principle on manipulated and nonmanipulated behaviors. *Journal of Applied Behavior Analysis*, *10*, 679–687.

Wolery, M., Kirk, K., & Gast, D. L. (1985). Stereotypic behavior as a reinforcer: Effects and side effects. *Journal of Autism and Developmental Disorders*, *15*, 149–161.

Received November 19, 2009
Final acceptance May 24, 2010
Action Editor, David Wacker