

*RESPONSE INTERRUPTION AND REDIRECTION FOR
VOCAL STEREOTYPY IN CHILDREN WITH AUTISM:
A SYSTEMATIC REPLICATION*

MEGAN DUFFY CASSELLA AND TINA M. SIDENER

CALDWELL COLLEGE

DAVID W. SIDENER

GARDEN ACADEMY

AND

PATRICK R. PROGAR

CALDWELL COLLEGE

This study systematically replicated and extended previous research on response interruption and redirection (RIRD) by assessing instructed responses of a different topography than the target behavior, percentage of session spent in treatment, generalization of behavior reduction, and social validity of the intervention. Results showed that RIRD produced substantial decreases in vocal stereotypy. Limitations of this study were that behavior reduction did not generalize to novel settings or with novel instructors and that appropriate vocalizations did not improve.

Key words: stereotypy, response interruption and redirection, autism

Response interruption has been shown to be effective in reducing various topographies of stereotypy (e.g., Azrin & Wesolowski, 1980; Sprague, Holland, & Thomas, 1997). Several studies have evaluated a variation of this intervention called response interruption and redirection (RIRD), which consists of interrupting the target response and redirecting the individual to engage in a different response (e.g., Ahearn, Clark, MacDonald, & Chung, 2007; Athens, Vollmer, Sloman, & St. Peter Pipkin, 2008). Ahearn et al. evaluated an RIRD

intervention in which the participants were interrupted and redirected to appropriate vocalizations via social questions and vocal imitation trials.

The purpose of the present study was to systematically replicate and extend Ahearn et al. (2007) by evaluating the effects of RIRD using directions that required motor responses on the vocal stereotypy of two children with autism, based on Ahearn et al.'s speculation that nonvocal tasks may produce similar treatment effects. In addition, motor responses may be manually prompted if necessary (whereas vocal responses may be difficult to prompt). Although we have noted that clinicians use this variation of RIRD with children with autism, few studies have evaluated it (e.g., Ahrens, Lerman, Kodak, Worsdell, & Keegan, 2011). The current study also extended previous research on RIRD for vocal stereotypy by including measurement of time spent in treatment, procedures to facilitate and assess the generalization of behavior reduction, and assessment of social validity.

This study is based on a thesis submitted by the first author, under the supervision of the second author, to the Department of Psychology at Caldwell College for the MA in Applied Behavior Analysis. We thank Christina Romeo, Kristie Murano, Sidona Anderson, Stacey Bord, Kim Sancho, Jill Bernstein, Nathan Cassella, Kristin Paladino, Allison Buleza, and Genevieve Williams for their assistance.

Address correspondence to Tina M. Sidener, Department of Applied Behavior Analysis, Caldwell College, 120 Bloomfield Ave., Caldwell, New Jersey 07006 (e-mail: tsidener@caldwell.edu).

doi: 10.1901/jaba.2011.44-169

METHOD

Participants and Setting

Participants were two boys who had been diagnosed with severe autism. Chris (7 years 2 months old) and Adam (4 years 11 months old) engaged in high levels of vocal stereotypy that occurred across the day and interfered with learning. Vocal stereotypy consisted of a variety of vocalizations and repetition of words and phrases. Prior to the study, both participants demonstrated that they could follow at least 15 one-step directions (e.g., stand up, turn around) and could imitate similar motor movements with 100% accuracy when modeled by an instructor. Sessions were conducted 5 days per week in three areas in the participants' school.

Three functional assessments were used to collect information about the possible function of the vocal stereotypy of the participants: the Functional Assessment Interview (FAI; O'Neill *et al.*, 1997), the Functional Assessment Screening Tool (FAST; Iwata & DeLeon, 1995) and the Motivation Assessment Scale (MAS; Durand & Crimmins, 1992). These assessments were conducted so that individuals who engaged in stereotypy with social functions could be excluded from the study. For Chris, responses to the FAI suggested an automatic reinforcement function, and rating scale results suggested automatic reinforcement or attention functions (MAS sensory $M = 6.25$; FAST attention proportional score = 0.75). For Adam, all assessments suggested an automatic reinforcement function (FAI; MAS sensory $M = 5$; FAST sensory proportional score = 0.83).

Response Measurement and Reliability

Vocal stereotypy was defined as noncontextual vocalizations, as well as contextual vocalizations repeated within 3 s of a similar vocalization (e.g., saying "ball" repetitively when seeing a ball). *Appropriate vocalizations* were defined as contextual vocalizations that were not repeated within 3 s of a similar vocalization. Time sampling was selected as the data-collection

system because it has been shown to result in fewer errors than either partial- or whole-interval recording (Alvero, Struss, & Rappaport, 2008). At the time of each measurement (i.e., at 10 s), the experimenter observed the participant for 5 s to obtain an adequate sample of the vocal stereotypy. A plus was scored if vocal stereotypy occurred at any time during the 5-s observation period. Observers used a 5-s observation window rather than a 1-s observation window so that they would be able to detect instances of repetitive vocalizations. During treatment sessions, data were also collected on time spent in treatment. These data were summarized as percentage of session in treatment by dividing total duration of RIRD by total session duration and multiplying by 100%.

Interobserver agreement on the occurrence of vocal stereotypy and appropriate vocalizations was assessed using the point-by-point method. An agreement was scored in an interval if both observers recorded the occurrence or nonoccurrence of vocal stereotypy or appropriate vocalizations. Agreement data were collected by a second independent observer for a mean of 43% of sessions. For vocal stereotypy, mean agreement was 93% (range, 84% to 100%) for Chris and 95% (range, 89% to 100%) for Adam. Interobserver agreement for appropriate vocalizations was 100% for both participants.

Design and Procedure

An ABAB reversal design was used to evaluate the effects of RIRD on vocal stereotypy. Baseline sessions were 5 min in duration, and treatment sessions ended following 5 min without the participant engaging in vocal stereotypy or after 30 min, whichever came first. The participant had noncontingent access to two moderately preferred nonauditory toys, identified via a multiple-stimulus without replacement preference assessment (DeLeon & Iwata, 1996), during all sessions. If the participant directed an appropriate vocalization toward the experimenter (e.g., "hi"), the experimenter briefly acknowledged it (e.g., "hi").

During baseline, the experimenter did not present demands or provide programmed consequences for vocal stereotypy. During treatment, each instance of vocal stereotypy was followed by the experimenter saying the participant's name, establishing attention by gaining eye contact, and giving a one-step direction that did not require a vocal response (e.g., "touch head"). The experimenter provided a model of the behavior if the participant responded incorrectly or did not respond within 5 s of the direction. Gentle manual prompting with graduated guidance was used if the child did not imitate the model following 5 s of the model being presented. Behavior-specific praise was delivered following each completed direction, and the procedure was terminated following three consecutive correct directed responses (with or without prompts) without engagement in vocal stereotypy. Ten different one-step directions were delivered during treatment.

Generalization was facilitated by implementing the procedure in two different rooms and by two different therapists. Generalization was assessed by evaluating the procedure in a novel room with an instructor who had not previously implemented the procedure. Probes of novel locations and instructors were conducted without the implementation of RIRD. Implementation of the treatment with novel demands was assessed as well.

To evaluate social validity, the children's parents and a teacher were asked to view 5-min videotaped segments of the participants during treatment and to respond to questions from a modified version of the Treatment Evaluation Inventory—Short Form (Van Norman, 2006).

Procedural integrity data were collected during a mean of 74% of sessions on correct implementation of directions, prompts, and praise, as well as the termination of directions at the correct time. Mean procedural integrity was 99.7% (range, 98% to 100%) for Chris and 99.6% (range, 97% to 100%) for Adam. Mean interobserver agreement on procedural integrity

was 99.6% (range, 94% to 100%) for Chris and 99.8% (range, 97% to 100%) for Adam and was calculated for a mean of 60% of procedural integrity sessions.

RESULTS AND DISCUSSION

Figure 1 shows the percentage of intervals with vocal stereotypy and the percentage of time spent in treatment during the treatment analysis and generalization probes. Vocal stereotypy for both participants was high and variable during baseline. RIRD resulted in an immediate and substantial decrease in vocal stereotypy that remained relatively stable throughout the phase. For both participants, vocal stereotypy was lower in the initial session during the return to baseline and subsequently increased as sessions continued. Effects of treatment were replicated in the second RIRD phase. The high level of vocal stereotypy in Session 30 for Adam coincided with illness.

Chris spent much less time (i.e., an average of less than 50% of sessions) in treatment than Adam, who spent most of the time during his sessions in treatment (i.e., an average of 77% of sessions). This may have been due to the reinforcing properties of stereotypy or attention provided during RIRD (i.e., vocal directions, manual prompts).

Use of multiple exemplars did not facilitate generalization across instructors or settings during probes when RIRD was not in effect (Figure 1). Vocal stereotypy remained at baseline levels during most of these sessions. Probes of RIRD with novel demands resulted in levels of vocal stereotypy similar to those in other treatment sessions. For Chris, appropriate vocalizations (data not shown) occurred at low levels during generalization probes in baseline and increased slightly during the first RIRD phase ($M = 2\%$; range, 3% to 9%). Adam emitted appropriate vocalizations in only the first RIRD phase ($M = 0.5\%$; range, 0% to 3%).

The social validity results for Chris indicated that his caregivers found RIRD highly accept-

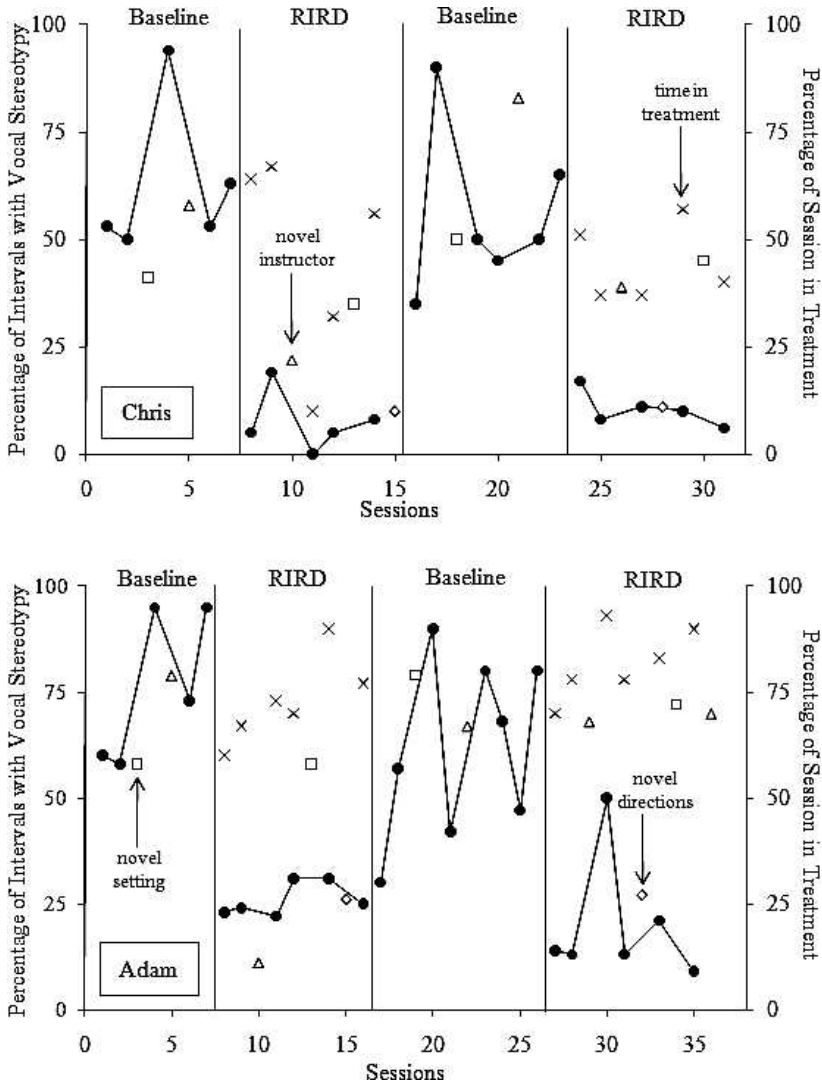


Figure 1. Percentage of intervals with vocal stereotypy and percentage of session spent in treatment for Chris (top) and Adam (bottom).

able ($M = 4.5$, range, 4 to 5). Results were slightly less favorable ($M = 4.1$, range, 3 to 5) for Adam. Adam's parents gave lower scores on the questionnaire regarding the acceptability and effectiveness of this procedure, perhaps because, as reflected in the videos, Adam spent most of the time during his sessions in treatment.

These results replicate and extend the findings of Ahearn et al. (2007) and demonstrate that RIRD can be effective in decreasing vocal stereotypy without the use of incompat-

ible responses. Some results obtained in the current study, however, do not replicate the findings of Ahearn et al. Appropriate vocalizations did not increase as vocal stereotypy decreased. In addition, vocal stereotypy did not remain low without RIRD.

It should be noted that, although this procedure is called redirection, the mechanism by which RIRD produces its effects is currently unknown. As posited by Ahearn et al. (2007), it is possible that presenting demands contingent

on vocal stereotypy functioned as a positive punishment contingency (see also Ahrens et al., 2011). Although toys were not removed during treatment, it is also possible that the intervention functioned as negative punishment via interruption of toy play contingent on vocal stereotypy.

One limitation of the current study was that sessions were often lengthy, with the majority of sessions lasting for 30 min. Ahearn et al. (2007) reported that many treatment sessions were longer than 10 min, but that many averaged 6 min. A second limitation was that aggression (i.e., pinching) and crying were observed in several sessions for Adam. These behaviors decreased as RIRD continued to be implemented. A final limitation is that momentary time sampling may both over- and underestimate behavior.

Future research might evaluate the implementation of RIRD by caregivers and in more natural environments. In addition, future research might examine the behavioral mechanisms in effect for RIRD by evaluating RIRD with and without toys, the effects of different contingencies for compliance (e.g., no rewards, praise), and the effects of different types of prompts (e.g., manual prompts, modeling). In future studies, researchers could also evaluate methods to promote generalization and appropriate vocalizations (e.g., by increasing the number of locations and instructors used in treatment, reinforcing appropriate vocalizations during RIRD).

REFERENCES

- Ahearn, W. H., Clark, K. M., MacDonald, R. P. F., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis, 40*, 263–275.
- Ahrens, E. N., Lerman, D. C., Kodak, T., Worsdell, A. S., & Keegan, C. (2011). Further evaluation of response interruption and redirection as treatment for stereotypic behavior. *Journal of Applied Behavior Analysis, 44*, 95–108.
- Alvero, A. M., Struss, K., & Rappaport, E. (2008). Measuring safety performance: A comparison of whole, partial, and momentary time-sampling recording methods. *Journal of Organizational and Behavior Management, 27*, 1–28.
- Athens, E. S., Vollmer, T. R., Sloman, K. N., & Pipkin, C. S. (2008). An analysis of vocal stereotypy and therapist fading. *Journal of Applied Behavior Analysis, 41*, 291–297.
- Azrin, N. H., & Wesolowski, M. D. (1980). A reinforcement plus interruption method of eliminating behavioral stereotypy of profoundly retarded persons. *Behaviour Research & Therapy, 18*, 113–119.
- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519–533.
- Durand, V. M., & Crimmins, D. (1992). *Motivation assessment scale*. Topeka, KS: Monaco & Associates.
- Iwata, B. A., & DeLeon, I. G. (1995). *The functional analysis screening tool (FAST)*. Unpublished manuscript, University of Florida.
- O'Neill, R. E., Horner, R. H., Albin, R. W., Sprague, J. R., Storey, K., & Newton, J. S. (1997). *Functional assessment and program development for problem behavior: A practical handbook*. Belmont, CA: Wadsworth.
- 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.
- Van Norman, R. K. (2006). *The effects of functional communication training, choice making, and an adjusting work schedule on problem behavior maintained by negative reinforcement* (Unpublished doctoral dissertation). The Ohio State University, Columbus.

Received November 11, 2009

Final acceptance July 9, 2010

Action Editor, Jennifer Zarcone

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.