Behavioral Interventions
Behav. Intervent. **24**: 1–15 (2009)
Published online 5 January 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/bin.270

EFFECTS OF ADJUSTING DRO SCHEDULES ON THE REDUCTION OF STEREOTYPIC VOCALIZATIONS IN CHILDREN WITH AUTISM

Eric Rozenblat[1,2]*, John L. Brown[2], Ann K. Brown[2,3], Sharon A. Reeve[1] and Kenneth F. Reeve[1]

[1]Caldwell College, NJ, USA
[2]REED Academy, NJ, USA
[3]The Graduate School and University Center of the City University of New York, NY, USA

The reduction of stereotypic behavior is important for individuals with developmental disabilities because it may interfere with learning new skills. A common procedure used to reduce stereotypic behavior is differential reinforcement of other behavior (DRO). A DRO schedule is a procedure in which reinforcement is delivered given the absence of a target response for a period of time. Although DRO schedules have been shown to be effective in reducing a variety of stereotypic behavior, empirical studies have not yet determined the most effective way to thin the reinforcement schedule. The current study examined the extent to which the relationship between the inter-response time (IRT) and the DRO requirement affects the reduction of stereotypic behavior. Results of the current study showed the DRO requirement that was set to the 25th percentile of the IRT distribution was more effective in reducing stereotypic behavior when compared with a DRO requirement set to the 95th percentile of the IRT distribution.

Because engagement in stereotypic behavior may interfere with learning new skills, the reduction of stereotypic behavior is important for individuals with developmental disabilities (Lovaas, Litrownik, & Mann, 1971). During the last 25 years functional analysis has become a primary behavior analytic approach for determining the function of problem behavior and corresponding treatment procedures (Iwata, 1982; Iwata, Dorsey, Slifer, Bauman, & Richman, 1994). It is the selection of a treatment procedure that corresponds to the function of the problem behavior that may increase the likelihood of successful treatment. Although functional analysis has received much empirical attention over the past 25 years, less attention has focused on improvements to the treatments themselves. For example,

*Correspondence to: Eric Rozenblat, 40F Lexington Lane, West Milford, NJ 07480, USA.
E-mail: eric_rozenblat@yahoo.com

when differential reinforcement is selected as an appropriate treatment there is little empirical evidence to guide the behavior analyst in determining the parameters of the reinforcement schedule.

Differential reinforcement of other behavior (DRO) is a procedure in which reinforcement is delivered if a target response does not occur during a given period of time, thus involving a contingency between the absence of a response and the delivery of the reinforcer (Repp, Felce, & Barton, 1991; Thompson, Iwata, Hanley, Dozier, & Samaha, 2003). In this schedule, there is a zero probability of reinforcement given a target response but an increased probability of reinforcement given the absence of a target response.

Behavior analysts who use DRO schedules to reduce stereotypic responses often attempt to thin the DRO schedule by increasing the duration of the DRO requirement as stereotypic responses decrease. Increasing DRO response requirements has the effect of thinning the reinforcement schedule (Barker & Thyer, 2000; Hegel & Ferguson, 2000; Kahng, Iwata, DeLeon, & Wallace, 2000; LeBlanc, Hagopian, Maglieri, & Poling, 2002; LeBlanc, Hagopian, & Maglieri, 2000; Leitenberg, Burchard, Burchard, Fuller, & Lysaght, 1977). Increasing the duration of the DRO requirement is desirable for a number of reasons. First, schedule thinning can reduce the probability of satiation to reinforcers. Longer DRO requirements also make it more practical to operate DRO schedules in clinical environments (Poling & Ryan, 1982). In addition, with longer DRO requirements, reinforcement becomes increasingly intermittent and consequently the reinforced responses should become more resistant to extinction (Shull & Grimes, 2006). Finally, thinning the reinforcement schedule may allow control of the response to shift from the DRO schedule to contingencies in the natural environment, thus achieving long-term maintenance of behavior reduction without continuing DRO schedules of reinforcement.

The most commonly addressed DRO schedule parameter is duration of the other behavior that is required to earn reinforcement during the initial implementation of the schedule (initial requirement). Additionally, four other parameters are relevant in studies that include thinning the schedule. The first of these is how much to increase the duration of the DRO requirement each time the schedule is thinned. Secondly, one must decide when to increase the duration of the DRO requirement. The final two parameters become relevant if the schedule must be made richer because it was thinned too rapidly or other uncontrolled variables lead to increases in stereotypic behavior. When such decreasing of the duration of the DRO requirement is desired, one must decide how much to decrease the duration of the DRO requirement and when to decrease the duration of the DRO requirement. Thus, there are five variables that may determine the effectiveness of DRO schedules: Setting the initial requirement, increasing the requirement, decreasing the requirement, determining when to increase, and when to decrease the requirements.

DOI: 10.1002/bin

Although there is no commonly accepted rule for determining the initial requirement, Repp et al. (1991) suggest that the initial requirement should be based on the mean response rate during baseline. Poling, Miller, Nelson, and Ryan (1978) set the initial DRO requirement to 32% of the mean inter-response time (IRT) during baseline. On the other hand, Ringdahl et al. (2002) set the requirement to 100% of the mean IRT. Barker and Thyer (2000) set their requirement to 158% of the mean IRT during baseline. There does not seem to be any consensus regarding how to use measures of baseline performances to set initial durations of DRO schedules. Given this variety of initial DRO values, it is unclear which is likely to be most effective.

Several studies addressed the issue of how to increase the duration of DRO requirements when rates of responding decreased. Some studies, such as Repp and Slack (1977), used three different initial requirement sizes that were set arbitrarily but increased the requirement duration based on a geometric progression. On the other hand, LeBlanc et al. (2000) increased the requirement by 33–50% each time the DRO value met criterion. Still, Thompson et al. (2003) increased the duration of the requirements by setting them equal to the mean IRT of the last 4 sessions, resulting in a self-adjusting DRO. Thus, with a variety of methods to increase the duration of the requirement, there is still a lack of empirical evidence indicating the most effective way to do so.

Some authors have suggested the conditions under which the duration of a DRO requirement should be increased. LeBlanc et al. (2000) set a rule that if response rates showed 90% or greater reduction from baseline for two consecutive sessions, the duration of the requirement should be increased. Thompson et al. (2003) implemented a rule that adjusted the DRO requirement every session so that it was equal to the mean IRT of the last four sessions. Other studies increased the duration of the requirement based on two successful intervals with no stereotypy (Ringdahl, Andelman, Kitsukawa, Winborn, Barretto, & Wacker, 2002). Peterson and Peterson (1968) reported attempts to increase the interval and Leitenberg et al. (1977) increased the duration of the requirement on a seemingly arbitrary basis. Thus, there are no clear empirical findings indicating when it is best to increase DRO requirements.

Few studies that decreased the duration of DRO requirements have been reported in the literature. LeBlanc et al. (2000) reduced the duration of the DRO requirement to the last successful requirement if responding did not meet criteria for advancement at the current requirement. Additionally, Thompson et al. (2003) recalculated the duration of the response requirement to equal the mean IRT from the previous four sessions, thus leading to decreasing of the DRO requirement if responding sufficiently increased. Aside from the two aforementioned studies, other literature reviewed in this paper did not address decreasing the duration of DRO requirements.

DOI: 10.1002/bin

LeBlanc et al. (2000) decreased the duration of the DRO requirement to the last successful requirement, if the response rates did not demonstrate 90% or greater reduction from baseline for two consecutive sessions. As previously mentioned, Thompson et al. (2003) decreased the duration of their DRO requirements every session based on the mean IRT of the last four sessions of that condition. Empirical examinations of decreasing the duration of DRO requirements are needed to guide behavior analysts when to adjust the DRO requirement and how often to adjust the DRO requirement. The purpose of the current study is to determine the extent to which the relationship between the distribution of IRT intervals and DRO intervals affects the levels of stereotypic vocalizations when using differential reinforcement schedules.

METHOD

Participants and Setting

Jeff, Joan, and Susan, ages 9, 9, and 10 respectively, participated in the study. All three learners were diagnosed with autism by an independent agency according to the criteria for autism as defined in the Diagnostic and Statistical Manual of Mental Disorders 4th ed., text revision (American Psychiatric Association, 2000). Prior to participation in this study, parents or legal guardians gave written consent for their children to participate. All children had prior experience with differential reinforcement schedules. For Jeff and Joan the study was conducted in their classrooms in a private behavior analytic school for children with autism. For Susan the study was conducted in an after-school home program during leisure activities.

Response Measurements

Episodes of stereotypic vocalizations were defined, as sounds that were not directed at another person, were not words, were not approximations of words, were words that were not appropriate in the current context, or words that were spoken below conversational loudness. An episode of stereotypic vocalizations began when the participant started engaging in stereotypic vocalizations. The end of an episode was defined as the participant not engaging in stereotypic vocalizations for 3 s. The duration of each episode of stereotypic vocalizations and DRO reinforcers were recorded using ABADat event recorder software (ABADat, 2007) on a Tungsten E2 Palm Pilot handheld computer. Data were summarized as the total duration of stereotypic vocalization episodes per session.

DOI: 10.1002/bin

Experimental Design

A reversal design consisting of a baseline condition, a 25th and 95th percentile condition were used. After an initial baseline phase, the 25th and 95th percentile conditions were introduced such that the first participant received the 25th percentile condition first, the second participant received the 95th percentile condition first, and the third participant received the 25th percentile condition first. For each of the three participants, the initial DRO requirement was determined from an analysis of the distribution of IRTs from the final baseline session. During each session, the DRO requirement was calculated based on the previous session of that condition. The adjustment of the DRO schedule for each participant was based on the participants' distribution of IRTs of stereotypic vocalizations. As stereotypic vocalizations varied from session to session, the distribution of IRTs also varied thus influencing the DRO requirement for the next session of that condition. For example, if during a session in the 25th percentile condition the IRT that corresponded to the 25th percentile was 15 s, during the next session of the 25th percentile condition the DRO requirement was set to equal 15 s. In some cases the DRO increased in time, thus thinning the schedule of reinforcement. In other cases the DRO requirement decreased in time, thus enriching the reinforcement schedule. Still, in other cases, the DRO requirement remained the same as in the previous session. If no stereotypic vocalizations occurred during a session, the DRO requirement was arbitrarily set to 60 min for the following session of that condition.

In a distribution of IRTs, each IRT was ranked from smallest to largest. The 25th percentile in the IRT distribution was the value that at least 25% of the IRTs were less than or equal to it, whereas the 95th percentile in the IRT distribution was the value that at least 95% of the IRTs were less than or equal to. Based on the current condition, the IRT that corresponded to either the 25th or 95th percentile was the DRO requirement for the next session of that condition.

Procedure

Each session was approximately 1 h in length. Sessions were conducted 2–3 times per day, as participants were available. During sessions, Jeff, Joan, and Susan were engaged in regularly scheduled activities.

At the beginning of each session, the experimenter reviewed the available reinforcers with the participant. Reinforcers were selected based on what teachers and parents thought the children would prefer. Additionally, if any of the participants requested other reinforcers, those reinforcers were made available as DRO reinforcers. If an episode of stereotypic vocalization occurred, the experimenter activated the event recorder. When the learner stopped engaging in stereotypic

DOI: 10.1002/bin

vocalizations for 3 s the observer deactivated the event recorder to mark the end of the episode. At the end of each episode of stereotypic vocalization, the DRO timer was restarted. If the DRO requirement elapsed and no episodes of stereotypic vocalization occurred, verbal praise, such as 'Great job working quietly,' and a reinforcer that the participant selected were delivered. During the time the participant was engaged with a DRO reinforcer, the DRO timer continued to run. For example, if the DRO requirement was set to 10 s and the participants did not engage in stereotypic vocalizations, a reinforcer was delivered. While the participants were consuming a reinforcer, the DRO timer continued running and participants could continue earning reinforcers each time the DRO requirement elapsed if they did not engage in episodes of stereotypic vocalization. If at any time during the DRO requirement the participants engaged in episodes of stereotypic vocalization, any activity DRO reinforcer was terminated and any edible reinforcer that was not in the participants' mouth was removed and the participant was returned to the regularly scheduled activities. Once the episode of stereotypic vocalization ceased, the DRO timer was restarted and participants could earn reinforcers when they met the DRO requirement.

Edible reinforcers that were delivered weighed approximately 1 g, which is approximately equal to one-half of a potato chip, one-quarter of an Oreo cookie, one m&m, or one Mike & Ike. Learners were allowed to engage in preferred activity reinforcers for 2 min.

Stereotypic Vocalization Interobserver Agreement (IOA)

The experimenter and another observer independently scored stereotypic vocalization. The primary experimenter trained five secondary observers by observing a learner who engaged in episodes of stereotypic vocalization. Training sessions varied between 15 and 60 min in length. Training continued for each secondary observer until he or she scored at least 80% agreement on vocal stereotypy. Agreements and disagreements of episodes of vocal stereotypy were scored on a point-by-point basis in 1 s bins. For an agreement to be scored, both the experimenter and the secondary observer had to score either a vocalization or a non-vocalization in a 1 s bin. A disagreement was scored if one observer scored a vocalization while the other observer scored a non-vocalization in a 1 s bin. Interobserver agreement (IOA) was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100.

For Susan, IOA data were collected for 40, 38, and 33% during the baseline, 25th percentile, and 95th percentile conditions respectively. During baseline the mean IOA was 91% for episodes of stereotypic vocalizations with a range of 87–95%. During the 25th and 95th percentile conditions, the mean IOA was 96 and 87%, respectively with

DOI: 10.1002/bin

a range of 93–98% during the 25th percentile condition and a range of 86–90% during the 95th percentile condition.

During baseline, IOA data were collected for Joan during 40% of the sessions. For the 25th and 95th percentile conditions, IOA data were collected for 36 and 43% of the sessions respectively. During baseline, the mean IOA for Joan's vocal stereotypy was 93% with a range of 91–94%. During the 25th and 95th percentile conditions the mean IOA was 99 and 97% during each respective condition. IOA ranged from 99 to 100% and 95 to 100% in each respective condition for Joan.

Finally, for Jeff, IOA data were collected for 29% of the sessions during baseline and for 35 and 33% of the sessions during the 25th and 95th percentile conditions. The mean IOA was 94% for episodes of vocal stereotypy during baseline with a range of 92–96%. During the 25th and 95th percentile conditions, IOA on Jeff's vocal stereotypy was 97% for both the 25th and 95th percentile conditions with ranges of 99–100% and 95–100% during each respective condition.

Procedural Integrity

For Susan, Joan, and Jeff, DRO reinforcer implementation accuracy was measured throughout each condition. The same scoring criterion and rules for scoring episodes of stereotypic vocalizations applied to scoring the implementation accuracy of the DRO schedule of reinforcement.

For Susan, procedural integrity data were 100% during baseline, and 97% respectively during the 25th and 95th percentile conditions. These data were collected for 40, 38, and 43% of the sessions respectively during the baseline, 25th and 95th percentile conditions. For Joan, data were collected during 40, 36, and 43% of the sessions during baseline, the 25th and the 95th percentile condition respectively. Procedural integrity data during each respective condition were 100, 99, and 96%. Finally, for Jeff procedural integrity data were collected for 29% of the sessions during baseline and accuracy was 100%. During the 25th and 95th percentile conditions, data were collected 35 and 33% of the sessions respectively with accuracy of 98 and 99% respectively.

Procedural Integrity IOA

The same scoring criterion and rules for scoring episodes of stereotypic vocalizations were applied to scoring the accuracy of the implementation of the DRO schedule of reinforcement. Observers recorded when reinforcement was delivered throughout the session in 1 s bins. Stereotypic vocalization data were used to create an optimal reinforcement schedule that indicated which 1 s bins should have been reinforced during each session. Accuracy in delivering DRO reinforcement was

DOI: 10.1002/bin

evaluated by comparing, on a point-by-point basis, the actual DRO reinforcer delivery to the optimal DRO reinforcer delivery. IOA was computed by dividing the number of bins when reinforcement was programmed and delivered plus the number of bins when no reinforcement was programmed and no reinforcement was delivered by the total number of bins in the session.

During baseline, the 25th and the 95th percentile condition, procedural integrity IOA data were collected for 40, 38, and 33% of the sessions respectively for Susan. For Joan, procedural integrity IOA data were collected for 40% of the baseline sessions, 36% of the 25th percentile condition sessions, and 43% of the 95th percentile condition sessions. Finally, procedural integrity IOA data for Jeff were collected for 29% of the sessions during baseline, 35% during the 25th percentile condition, and 33% of the sessions during the 95th percentile condition.

For all three participants, procedural integrity IOA during baseline was 100%. For Susan, during the 25th and 95th percentile conditions IOA was 97 and 96% respectively with ranges of 96–98% and 87–100%. For Joan, IOA was 99 and 100% in the 25th and 95th percentile condition respectively with a range of 99–100% during the 25th percentile condition and 100% during the 95th percentile condition. Finally, IOA was 99 and 100% during the 25th and 95th percentile conditions respectively for Jeff with a range of 98–100% during the 25th percentile condition and 100% during the 95th percentile condition.

RESULTS

Figure 1 shows the total duration of episodes of stereotypic vocalizations in each session for Susan. During baseline, Susan engaged in episodes of stereotypic vocalizations for a mean of 34 min. During the first 95th percentile condition Susan spent similar amounts of time engaging in episodes of stereotypic vocalizations as compared to baseline sessions. The mean duration of episodes of stereotypic vocalizations during this condition was 31 min. Following this condition, the 25th percentile condition was introduced. There was an immediate decrease in the total duration of episodes of stereotypic vocalizations. During this condition Susan did not engage in episodes of stereotypic vocalizations for more than 6 min with the exception of one session. During that session, Susan spent less than 12 min engaging in episodes of stereotypic vocalizations. The mean during this condition was 7 min.

When the 95th percentile condition was reintroduced for Susan, episodes of stereotypic vocalizations immediately increased. With the exception of one data point, durations of episodes of stereotypic vocalizations were between 40 and 51 min. The mean during this condition was 37 min. Finally, during the last 25th percentile condition, data showed similar levels of stereotypic vocalization to those reached

DOI: 10.1002/bin

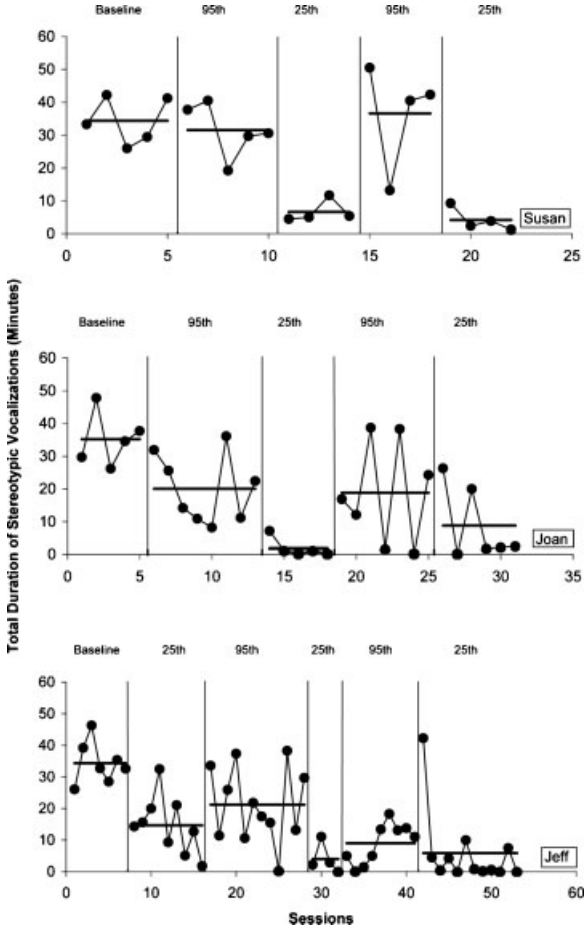

Figure 1. Total duration of episodes of stereotypic vocalizations across two DRO requirements for Joan, Susan, and Jeff. Horizontal lines in each phase represent the mean for that phase.

during the first 25th percentile condition. Throughout this condition, Susan spent less than 10 min engaging in stereotypic vocalizations in each session. During the final three sessions, Susan did not spend more than 4 min engaging in stereotypic vocalizations. The mean of duration of stereotypic vocalizations was 4 min, the lowest mean of all the conditions.

During baseline Joan engaged in stereotypic vocalizations for similar amounts of time as Susan did during her baseline condition. When the 95th percentile condition was introduced to Joan, the mean amount of time she spent engaging in stereotypic vocalizations was 20 min. Following the 95th percentile condition, the 25th percentile

DOI: 10.1002/bin

condition was introduced. During this condition, the mean amount of time Joan spent engaging in stereotypic vocalizations was 2 min. The subsequent two conditions were similar to one another with the exception of a higher mean of 9 min during the final 25th percentile condition.

For Jeff baseline durations of stereotypic vocalizations were similar to that of Susan and Joan. Additionally, both the 25th and the 95th percentile conditions were similar to the other participants in that lower means of stereotypic vocalizations were observed in the 25th percentile condition as compared to the 95th percentile condition.

Table 1 shows the programmed DRO requirement in seconds for each participant during consecutive sessions for each condition. During the 25th percentile condition Susan's, Joan's, and Jeff's DRO requirements ranged from 2 to 22 s, 1 to 3600 s, and 1 to 3600 s respectively. During the 95th percentile sessions Susan's, Joan's, and Jeff's DRO requirements ranged from 17 to144 s, 35 to 3591 s, and 26 to 3803 s respectively.

Table 1. Consecutive DRO requirements in seconds within conditions for each participant.

	Susan		*Joan*		*Jeff*	
Session	*25th*	*95th*	*25th*	*95th*	*25th*	*95th*
1	2	18	2	63	4	65
2	22	23	1	44	5	58
3	13	18	35	101	6	453
4	6	61	8	195	2	81
5	19	41	3	83	9	49
6	19	34	360	317	2	190
7	8	17	3	100	12	84
8	16	144	3600	66	3	120
9		44	3	79	13	96
10			3600	65	5	3600
11			4	64	1301	52
12			3	78	7	212
13			13	1375	1	127
14				35	3600	63
15				3591	6	42
16					1	81
17					4	744
18					1	222
19					3600	137
20					10	127
21					3	80
22					19	
23					2	
24					3600	
25					9	

DOI: 10.1002/bin

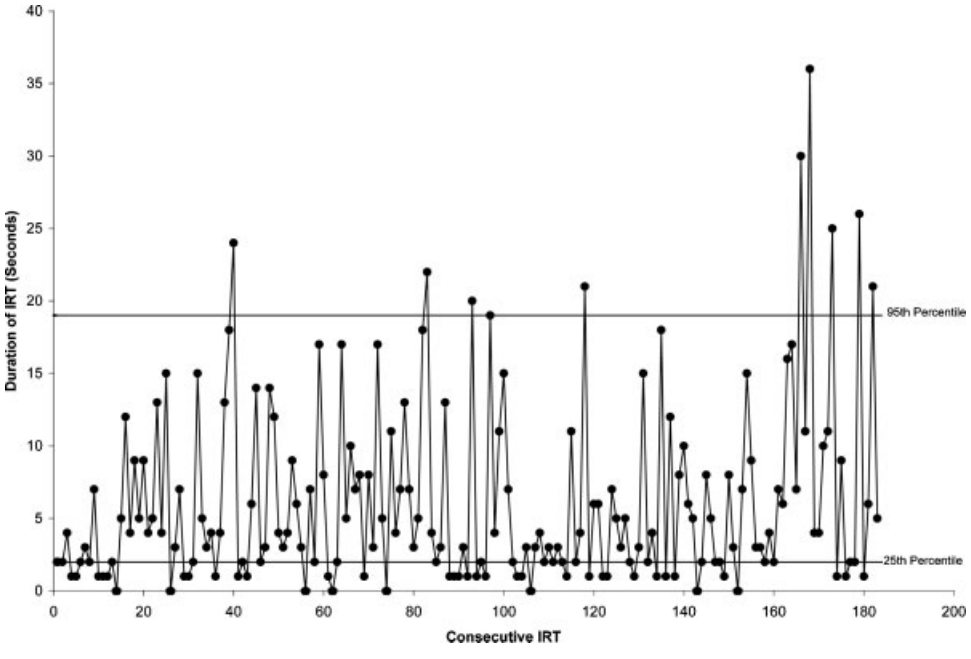

Figure 2. Duration of consecutive inter-response times for Susan during Session 5. The horizontal lines represent the 25th percentile rank and the 95th percentile rank during the session.

Figure 2 graphically shows how the 95th and 25th percentile IRTs were identified. This figure shows the duration of each consecutive IRT for Susan during session five. The horizontal line marked 95th Percentile indicates the level at which 95% of the inter-response fall below. Similarly, the horizontal line marked 25th Percentile indicates the level at which 20% of the IRTs fall below.

DISCUSSION

In this study a reversal design demonstrated experimental control of stereotypic vocalizations. Overall, the 25th percentile was more effective in decreasing episodes of stereotypic vocalization as compared to baseline and the 95th percentile conditions. The 95th percentile condition was not as effective in decreasing episodes of stereotypic vocalizations when compared to the 25th percentile condition but was effective when compared to the baseline phase of the study for Joan and Jeff. For Susan, the 95th percentile and baseline conditions were similar in the duration of stereotypic vocalizations.

DOI: 10.1002/bin

The effects of two different DRO conditions in reducing episodes of stereotypic vocalizations were compared. As stereotypic responding changed, DRO requirements were recalculated and set to the duration of the corresponding IRT percentile from the previous session of that condition. The current study differed from previous research in that frequently adjusting DRO schedules were used. These schedules reflected each participant's distribution of IRTs from the previous session of that condition. After each session in each condition, a DRO value was recalculated and the DRO requirement for the subsequent session was adjusted to reflect each participant's rate of responding from the previous session of that condition.

This study used DRO requirements that were based on the IRT distributions to maximize the effectiveness of the DRO schedule by setting DRO requirements based on the participant's IRTs from the prior session. DRO requirements varied in length from session to session with a range of 1–3600 s across conditions. During the 25th percentile condition, DRO requirements were relatively short and the participants were more likely to earn reinforcers and not engage in episodes of stereotypic vocalizations. During the 95th percentile condition, DRO requirements were greater and there were many instances in which participants did not earn any reinforcers during the session.

As described earlier, the literature indicated that LeBlanc et al. (2000) and Thompson et al. (2003) increased and decreased DRO requirements. LeBlanc et al. (2000) increased the requirements using a seemingly arbitrary criterion of 33–50% when responding was reduced by 90% or greater as compared to baseline and the DRO requirement was decreased to the last successful requirement when this criterion was not met. One way that the current study differed from LeBlanc et al. (2000) was that each session the DRO requirement increased or decreased based on the performance of the previous session of that condition and no pre-determined performance criterion was used. These DRO requirements adjusted according to each participant's performance during each session. Thompson et al. (2003) recalculated DRO requirements that were equal to the mean IRT from the previous four sessions. Adjusting a DRO requirement from the mean of four previous sessions could result in not achieving optimal effectiveness as the mean may not be the most relevant performance parameter and change in the reinforcement schedule is infrequent. For example, if the mean IRT of four sessions were 5, 6, 9, and 100 s, the mean IRT would be set to 30 s. In this case the outlying score dramatically increases the DRO requirement.

One weakness in the current study is that it was run in 1-hour sessions and those sessions could be separated by minutes, hours, or even days based on the availability of the participants. For example, if a one-hour session followed another session a couple of days later, the DRO requirement that was set based on the previous session could drastically affect the performance of the individual for a number of reasons.

DOI: 10.1002/bin

First, because there could be days in between data collection periods, a number of factors such as lack of sleep, sickness, or even weather could have affected how the participant responded. Thus, if the DRO requirement was set to 10 min based on the performance of the previous session, all of the preceding factors may have contributed to greater durations spent engaging in vocalizations that in turn could have lead to decreasing DRO requirements. This was a potential reason for the variability in the total duration of episodes of stereotypic vocalizations for each participant.

The goal of this procedure was to reduce stereotypy and increase the duration of the DRO requirement. For example, during a short DRO requirement of 20 s each participant could have earned the maximum number of reinforcers for that session and never engaged in stereotypic vocalizations. Following that session, if the participant never engaged in stereotypic vocalizations, by rule the DRO requirement was automatically increased to 60 min. This was an arbitrary rule determined prior to the beginning of the study. However, increasing the time from 20 s to 60 min was too big an increase and the thinning of the reinforcement schedule happened far to quickly and participants never earned reinforcers. Clinically, it may take several months and even years to have DRO requirements set to equal the entire length of a school day. So, in this study as DRO requirements rapidly increased, participants spent more time engaging in episodes of stereotypic vocalizations thus leading to smaller DRO requirements on the subsequent session which in turn led to participants spending less time engaging in episodes of stereotypic vocalizations. Moving DRO requirements very rapidly may not be as effective as moving them more systematically.

Another potential weakness is that collecting a 1-hour sample of data may not be long enough to accurately obtain sufficient information about how each participant would respond throughout a 6-hour school day. There are many factors in one school day that could affect how a learner responds. One factor could be certain teachers could be associated with the way in which an individual responds. For example, a learner may respond differently with different teachers such that a learner could prefer one teacher instead of another teacher because one teacher may do less preferred activities with the learner. Thus, how the learner responds to each teacher could affect his or her rate of stereotypy. For example, throughout a school day some teachers may provide a higher rate of reinforcement which may lead to learners spending less time engaging in stereotypy, whereas other teachers may not deliver enough reinforcement which could lead learners to spend more time engaging in stereotypy. Consequently, a 1-hour sample of data collection does not take into consideration how each participant would respond throughout an entire school day.

The clinical use of this procedure may be limited because it is difficult to operate. For example, during this study DRO requirements were as low as 1 s. To operate this procedure clinically, it may require an additional teacher if the DRO requirement is of

DOI: 10.1002/bin

short duration and it may be difficult for one teacher to teach, record, collect data, and deliver reinforcers if the DRO requirement is short.

When DRO requirements were low and participants were earning frequent reinforcers, it was hypothesized that satiation would occur. However, each participant varied his or her choices of reinforcers and satiation did not seem to occur for any of the participants.

Future research could investigate other ways in which to adjust DRO requirements as each individual learner's responding changes. As a learners' behavior within a session changes, the DRO requirement could change within that session to reflect that behavior. Self-adjusting DRO schedules within sessions could be more effective in reducing stereotypic behavior.

ACKNOWLEDGEMENTS

The authors thank Jeff, Joan, and Susan for all of their hard work on this project. They also thank their parents for supporting their participation. In addition the following individuals provided essential assistance with data collection: Ann Marie Agresti, Nicole Elmera, Leah Fanning, and Shannon Monaghan.

REFERENCES

ABADat. (2007). Event recorder software. Retrieved January 2, 2007, from http://www.abadat.com.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.

Barker, K. L., & Thyer, B. A. (2000). Differential reinforcement of other behavior in the treatment of inappropriate behavior and aggression in an adult with mental retardation at a vocational center. *Scandinavian Journal of Behaviour Therapy*, *29*, 37–42.

Hegel, M. T., & Ferguson, R. J. (2000). Differential reinforcement of other behavior (DRO) to reduce aggressive behavior following traumatic brain injury. *Behavior Modification*, *24*, 94–101.

Iwata, B. A. (1982). Toward a functional analysis of self-injury. *Analysis & Intervention in Developmental Disabilities*, *2*, 3–20.

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.

Kahng, S., Iwata, B. A., DeLeon, I. G., & Wallace, M. D. (2000). A comparison of procedures for programming noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis*, *33*, 223–231.

LeBlanc, L. A., Hagopian, L. P., & Maglieri, K. A. (2000). Use of a token economy to eliminate excessive inappropriate social behavior in an adult with developmental disabilities. *Behavioral Interventions*, *15*, 135–143.

DOI: 10.1002/bin

LeBlanc, L. A., Hagopian, L. P., Maglieri, K. A., & Poling, A. (2002). Decreasing the intensity of reinforcement-based interventions for reducing behavior: Conceptual issues and a proposed model for clinical practice. *The Behavior Analyst Today*, *3*, 289–300.

Leitenberg, H., Burchard, J. D., Burchard, S. N., Fuller, E. J., & Lysaght, T. V. (1977). Using positive reinforcement to suppress behavior: Some experimental comparisons with sibling conflict. *Behavior Therapy*, *8*, 168–182.

Lovaas, O. I., Litrownik, A., & Mann, R. (1971). Response latencies to auditory stimuli in autistic children engaged in self-stimulatory behavior. *Behaviour Research and Therapy*, *9*, 39–49.

Peterson, R. F., & Peterson, L. A. (1968). The use of positive reinforcement in the control of self-destructive behavior in a retarded boy. *Journal of Experimental Child Psychology*, *6*, 351–360.

Poling, A., Miller, K., Nelson, N., & Ryan, C. (1978). Reduction of undesired classroom behavior by systematically reinforcing the absence of such behavior. *Education and Treatment of Children*, *1*, 35–41.

Poling, A., & Ryan, C. (1982). Differential-reinforcement-of-other-behavior schedules. *Behavior Modification*, *6*, 3–21.

Repp, A. C., Felce, D. L., & Barton, L. E. (1991). The effects of initial interval size on the efficacy of DRO schedules of reinforcement. *Exceptional Children*, *57*, 417–425.

Repp, A. C., & Slack, D. J. (1977). Reducing responding of retarded persons by DRO schedules following a history of low-rate responding: A comparison of ascending interval sizes. *The Psychological Record*, *27*, 581–588.

Ringdahl, J. E., Andelman, M. S., Kitsukawa, K., Winborn, L. C., Barretto, A., & Wacker, D. P. (2002). *Behavioral Interventions*, *17*, 43–49.

Shull, R. L., & Grimes, J. A. (2006). Resistance to extinction following variable-interval reinforcement: Reinforcer rate and amount. *Journal of the Experimental Analysis of Behavior*, *85*, 23–39.

Thompson, R. H., Iwata, B. A., Hanley, G. P., Dozier, C. L., & Samaha, A. L. (2003). *Journal of Applied Behavior Analysis*, *36*, 221–238.

DOI: 10.1002/bin