A comparison of differential reinforcement procedures for treating automatically reinforced behavior

CHELSEA B. HEDQUIST AND EILEEN M. ROSCOE
WESTERN NEW ENGLAND UNIVERSITY, THE NEW ENGLAND CENTER FOR CHILDREN

Children diagnosed with autism spectrum disorder (ASD) often exhibit stereotypy, which can be socially stigmatizing, interfere with daily living skills, and affect skill acquisition. We compared differential reinforcement of alternative behavior (DRA) and differential reinforcement of other behavior (DRO) when neither procedure included response blocking or interruption for (a) reducing stereotypy, (b) increasing task engagement, and (c) increasing task completion. DRA contingencies yielded superior outcomes across each measure when evaluated with 3 individuals with autism spectrum disorder.

Key words: automatically reinforced behavior, autism spectrum disorder, DRA, DRO, stereotypy

Stereotypy, defined as repetitive motor movements and vocalizations that serve no adaptive purpose (Koegel, Firestone, Kramme, & Dunlap, 1974), is often exhibited by individuals diagnosed with intellectual disabilities, including autism spectrum disorder (ASD; Bodfish, Symons, Parker, & Lewis, 2000; Matson & Dempsey, 2008). Stereotypy may be socially stigmatizing (Jones, Wint, & Ellis, 1990); limit opportunities for appropriate social interaction; interfere with daily living skills such as leisure engagement, vocational tasks, and other self-care tasks (Dunlap, Dyer, & Koegel, 1983; Koegel & Covert, 1972; Koegel et al., 1974; Lanovaz, Robertson, Soerono, & Watkins, 2013; Rapp & Vollmer, 2005); and interfere with skill acquisition in academic settings (Lovaas, Litrownik, & Mann, 1971). Due to these concerns, stereotypy is a common target for assessment and intervention.

According to Rapp and Vollmer (2005), more than 90% of published functional analyses of stereotypy suggested maintenance by automatic reinforcement. Automatically reinforced problem behavior poses a challenge to effective treatment because the specific reinforcer cannot be readily identified, and as a result, the contingency between behavior and its reinforcer cannot be eliminated via extinction as it could with socially reinforced problem behavior.

Despite this challenge, differential reinforcement of other behavior (DRO) and differential reinforcement of alternative behavior (DRA) have been used to effectively treat stereotypy. DRO procedures involve delivering a putative reinforcer following the abstention of a behavior for a predetermined interval (Catania, 2013); this contingency has reduced automatically reinforced skin picking (Tiger, Fisher, & Bouxsein, 2009), vocal stereotypy (Taylor, Hoch, & Weissman, 2005), and motor ticks (Capriotti, Brandt, Ricketts, Espil, & Woods, 2012). For example, Taylor et al. (2005) decreased a participant’s vocal stereotypy by arranging a 1-min DRO. After attaining initial reductions in vocal stereotypy, the DRO interval was increased to 5 min while sustaining low levels of stereotypy.

DRA procedures involve delivering a putative reinforcer following the occurrence of a specific alternative behavior. For example,
Potter, Hanley, Augustine, Clay, and Phelps (2013) decreased stereotypy by arranging DRA contingencies to support item engagement with three individuals with ASD. This procedure specifically involved manually disrupting instances of stereotypy and allowing access to stereotypy contingent on item engagement. Similar to Potter et al., most DRA demonstrations with stereotypy as the reinforcer include response blocking or response interruption (e.g., Hanley, Iwata, Thompson, & Lindberg, 2000) to minimize the extent to which stereotypy continues to produce automatic reinforcement. However, the inclusion of response blocking or interruption may limit the utility of these procedures in practical application.

Response blocking and interruption are labor intensive in that they require a dedicated caregiver to disrupt each instance of stereotypy. The practical costs of staffing such a program may prohibit adoption of the procedure in most settings. Additionally, the introduction of marginally intermittent or delayed schedules for response blocking and interruption has failed to maintain behavioral reductions (Kleibert & Tiger, 2011; Smith, Russo, & Duy, 1999). Even with dedicated staffing, high procedural integrity for response blocking and interruption of each instance of stereotypy is unlikely. In light of these limitations associated with response blocking and interruption, an evaluation of DRA and DRO without response blocking and interruption is warranted. However, the utility of DRA alone in the treatment of stereotypy is not well established. In one published example, Vollmer, Marcus, and LeBlanc (1994) examined DRA without response blocking or interruption with two of three individuals with automatically reinforced stereotypic hand mouthing and self-injury. During DRA, therapists activated a sound-producing leisure item contingent on hand–item contact. The DRA procedure alone decreased self-injury for one participant, whereas the addition of a blocking and hands down procedure was necessary to decrease hand mouthing with the other participant.

The purpose of the current study was to compare DRO and DRA contingencies in the treatment of stereotypy when neither procedure included response blocking or response interruption. In addition, caregiver-nominated vocational activities from each participant’s Individualized Education Program (IEP) were included during both DRO and DRA evaluations, permitting comparison across problem behavior, activity engagement, and productivity.

METHOD

Participants and Settings
Three individuals diagnosed with ASD, who attended a residential school for children with intellectual disabilities, participated. Each was referred for participation by their clinical teams due to stereotypy that interfered with daily activities. John and Nick were 14-year-old young men, and Solomon was a 16-year-old young man. Each student communicated vocally in short phrases and demonstrated receptive identification, imitation, matching, and limited computer skills.

Sessions were conducted in the participants’ classroom while no other students were present (John), or in a small research room (Nick and Solomon). Each room had a chair, table, cabinet, and task materials.

Response Measurement and Interobserver Agreement
A primary data collector scored stereotypy, task engagement, and task productivity during each session. Stereotypy, scored via continuous duration recording, was defined as any repetitive, non-contextual, motor movement. Specific forms of stereotypy for John included hand wringing, circular waving, finger swiping on desk, and other repetitive hand movements. For Nick, stereotypy included body rocking, hand waving, hand flapping, finger tapping, and finger flicking. For Solomon, stereotypy included body rocking,
hand flapping, and skin picking. Engagement, scored via continuous duration recording, was defined as any hand-to-item contact with task materials in the absence of stereotypy. Stereotypy and engagement were considered mutually exclusive; if the participant engaged in item contact while engaging in stereotypy, observers scored only stereotypy. Productivity was measured by counting the number of tasks completed. For John and Solomon, a product was counted following a sequence of folding the item, placing the item in a pile, and removing his hand. For Nick a product was counted after he stamped a notecard, put the notecard in an envelope, placed the envelope in a completed pile, and removed his hands.

To assess interobserver agreement (IOA), a second observer scored video records using the same observational codes for 33% of sessions. The first author calculated IOA for stereotypy and engagement using mean duration-per-occurrence within 10-s intervals. Observers’ records were compared, the shorter duration was divided by the longer duration for each interval, the quotients were averaged across the intervals and then multiplied by 100. The first author calculated mean count-per-interval IOA for productivity by comparing observers’ records, dividing the smaller number of responses by the larger number in each 10-s interval, averaging the quotients and multiplying by 100. For John, mean IOA was 81% (range, 76% to 96%) for stereotypy, 84% (range, 75% to 97%) for engagement, and 81% (range, 68% to 98%) for productivity. For Nick, mean IOA was 92% (range, 80% to 100%) for stereotypy, 91% (range, 81% to 100%) for engagement, and 91% (range, 83% to 100%) for productivity. For Solomon, mean IOA yielded 90% (range, 85% to 100%) for stereotypy, 91% (range, 84% to 100%) for engagement, and 92% (range, 86% to 100%) for productivity.

Functional Analysis
A functional analysis was conducted to rule out social sources of reinforcement as maintaining variables for stereotypy. During the first phase, the first author conducted a series of 10-min alone sessions (Querim et al., 2013), in which the participant was alone in the room such that no social consequences were provided following stereotypy. During the second phase, 10-min alone, attention, demand, and play conditions were alternated using procedures similar to Iwata, Dorsey, Slifer, Bauman, and Richman (1994/1982). Figure 1 depicts the functional analysis results for John (top panel), Nick (middle panel), and Solomon (bottom panel). When exposed to only alone sessions, participants exhibited moderate to high levels of stereotypy. Because Solomon exhibited inconsistent levels of stereotypy during the first 11 sessions, materials typically present in his classroom (e.g., vocational stimuli, match-to-sample stimuli, silverware, etc.) were added for the remainder of the analysis to create a more ecologically valid environment. Solomon’s stereotypy increased during the final three sessions of the alone condition. During the multielement phase, when all conditions were conducted, stereotypy occurred at moderate-to-high levels across conditions, with slightly lower levels in the attention (all participants) and play (Solomon) conditions. These results indicated that stereotypy was not socially maintained as no social-test condition resulted in increased levels (i.e., a reinforcement effect).

Edible Preference Assessment
A paired-stimulus preference assessment (PSPA; Fisher et al., 1992) was conducted to identify three highly preferred edibles to use during subsequent treatment sessions. Prior to each treatment session, the therapist presented these three edibles in an array, instructed the participant to “pick one” and delivered the selected edible according to the DRO and DRA contingencies described below.
A task assessment was conducted to identify a previously mastered vocational task from each participant’s IEP that he could complete independently. For each participant, five to six tasks were included in the task assessment. These included folding shirts and towels (John and Nick), stamping and stuffing envelopes (Nick), stuffing envelopes (Solomon), tracing letters (John), sorting silverware (John, Nick, and Solomon), stapling (Nick), hole punching (Nick), silverware wrapping (Solomon), and sorting office supplies (Solomon). The task assessment consisted of two phases. The first phase included sessions of 10 trials. During each trial, the therapist presented an instruction to complete the target task and delivered a preferred edible reinforcer contingent on completion. Tasks that were completed independently on 80% or more trials were included in the subsequent phase. During the second phase, the therapist presented each task, singly, for 5 min and provided an initial instruction to engage with the materials. There were no further prompts or reinforcers provided for completing the task. The goal of this assessment was to select a task that participants (a) completed when prompting and edible reinforcement were provided, (b) did not perform consistently in the absence of prompting and edible reinforcement. Based on the results of this assessment, shirt and towel folding was selected for John (8% engagement) and Solomon (32% engagement), and stamping and stuffing envelopes was selected for Nick (28% engagement).

**Treatment Assessment**

During the treatment assessment, conditions included a no-reinforcement baseline, DRA, and DRO. The effects of DRA and DRO were evaluated using a combination of multielement and reversal designs. The relative and absolute effects of DRA and DRO were evaluated using a multielement design, and the absolute effects of DRA and DRO relative to baseline were evaluated using a reversal design. Three to five 5-min sessions were conducted per observation period, and three to six observation periods were conducted per week.

Before each observation period, the therapist presented the three edibles selected from the PSPA and vocally prompted the participant to, “pick one.” The selected item was used for all
sessions within that observation period. At the beginning of each session, the therapist sat next to the participant, presented task materials, and stated, “You can do [task] if you want to.”

During baseline sessions, two different colored versions of the target task were presented, that would later be associated with DRO and DRA conditions (e.g., a blue set versus a white set of towels or a green versus pink poster board under task materials). The first author drew slips of paper from a hat to determine the alternation pattern of DRA and DRO sessions, with the expectation that no more than two consecutive DRA or DRO sessions were allowed. There were no social consequences for motor stereotypy or productivity during baseline sessions.

The DRA sessions were similar to baseline except the therapist placed edibles within the participant’s eyesight and delivered one edible directly into the participant’s mouth for each completed task. There were no programmed consequences in effect for motor stereotypy.

The DRO sessions were similar to baseline except that the therapist delivered one edible directly into the participant’s mouth at the end of each omission interval. The first author identified omission intervals by calculating 85% of the mean interresponse time between bouts of stereotypy during the last three baseline sessions for each participant. This calculation yielded omission intervals of 30 s (John), 14 s (Nick), and 7 s (Solomon). If stereotypy occurred before the interval ended, the therapist reset the timer without comment and did not respond to task engagement at any point during sessions.

RESULTS

Figure 2 depicts the results from John’s treatment analysis. During baseline, stereotypy, task engagement, and productivity occurred at moderate and variable levels, with no tasks completed and no systematic differences across material sets (i.e., white and blue task materials to be associated with DRO and DRA). When DRO and DRA were implemented in a multielement comparison, stereotypy decreased across conditions with minimal differences between DRO and DRA (M = 12.1% of session during DRO relative to 28.6% in baseline; M = 7.8% of session during DRA relative to 22.1% in baseline). The return to baseline produced small increases in stereotypy to near baseline levels (M = 20.6% of session for DRO and 20.5% of session for DRA). No systematic changes from baseline occurred in engagement or productivity across the first three phases of the treatment analysis. Because the initial multielement comparison of DRO and DRA may have been susceptible to carryover effects, each treatment was evaluated individually relative to baseline in subsequent phases. The DRA procedure decreased stereotypy (M = 6.8% of session relative to 27.9% in baseline), increased task engagement (M = 78.7% of session relative 45.3% in baseline), and increased productivity (M = 3.4 products per min [ppm] relative to 1.5 ppm in baseline). The DRO procedure did not result in a meaningful reduction in stereotypy (M = 14.1% of session relative to 19.9% in baseline); however, engagement and productivity increased slightly (M = 71.5% engagement relative to 56.4% in baseline; M = 2.5 ppm relative to 1.8 ppm in baseline). When the DRA versus baseline phase was reimplemented, stereotypy decreased (M = 2.1% of sessions relative to 49.0% in baseline), engagement increased (M = 94.9% of sessions relative to 28.8% in baseline), and productivity increased (M = 4.6 ppm relative to 1.2 ppm in baseline).

Figure 3 depicts the results of the treatment analysis for Nick. During baseline, stereotypy and task engagement occurred at moderate and variable levels, with no tasks completed and no systematic differences between material sets. Both the DRO and DRA contingencies produced decreased stereotypy (M = 1.6% of session in DRO relative to 38.4% in baseline; M = 1.1% of session in DRA relative to 33% in baseline),
increased engagement ($M = 90.1\%$ of session in DRO relative to $35.4\%$ in baseline; $M = 93.0\%$ of session in DRA relative to $28.2\%$ in baseline), and increased productivity ($M = 1.3$ ppm in DRO relative to $0.0$ in baseline; $M = 2.3$ ppm in DRA relative to $0.0$ in baseline). A return
Figure 3. For Nick, the top panel depicts motor stereotypy, the second engagement, and the third productivity across baseline (BL), differential reinforcement of alternative behavior (DRA) compared to differential reinforcement of other behavior (DRO), return to BL, DRA compared to BL, DRO compared to BL, then a return to DRA compared to BL. During baseline the only difference between the conditions was the color of the poster board under the materials.
to baseline produced increased stereotypy ($M = 34.8\%$ of session in DRO baseline relative to 45.6\% in DRA baseline), decreased task engagement ($M = 35.3\%$ of session in DRO baseline relative to 35.3\% in DRA baseline), and decreased productivity ($M = 0.3$ ppm in DRO baseline relative to 0.4 ppm in DRA baseline). There were no systematic differences between the DRA and DRO conditions across the first three phases. As with John, DRA and DRO were subsequently evaluated individually relative to baseline. DRA contingencies decreased stereotypy ($M = 0.5\%$ of session relative to 41.0\% in baseline), increased task engagement ($M = 95.4\%$ of session relative to 45.9\% in baseline), and increased productivity ($M = 4.2$ ppm relative to 0.0 in baseline). The DRO procedure did not result in reduced stereotypy ($M = 26.3\%$ of session relative to 20.7\% in baseline), nor did it increase task engagement ($M = 67.8\%$ of session relative to 0.0 in baseline). When DRA was reimplemented, stereotypy decreased ($M = 1.2\%$ of session relative to 75.5\% in baseline), engagement increased ($M = 96.3\%$ of session relative to 17.3\% in baseline), and productivity increased ($M = 4.5$ ppm relative to 0.0 in baseline). Next, the DRA schedule was faded by delivering tokens that could be exchanged for an edible and by gradually increasing the number of tokens required for an exchange from 1 to 2, to 4, and to 20.

Figure 4 depicts the results of the treatment analysis for Solomon. During baseline, stereotypy occurred at moderate to high levels, and task engagement and productivity occurred at low levels. No systematic difference between material sets was observed. In the DRA and DRO treatment comparison phase, similar outcomes occurred across conditions. Stereotypy decreased ($M = 18.1\%$ of session in DRO relative to 57.8\% in baseline; $M = 14.3\%$ of session in DRA relative to 63.0\% in baseline), engagement increased ($M = 53.3\%$ of session in DRO relative to 13.3\% in baseline; $M = 60\%$ of session in DRA relative to 15.7\% in baseline), and productivity increased ($M = 2.9$ ppm in DRO relative to 0.8 ppm in baseline; $M = 3.7$ ppm in DRA relative to 0.9 ppm in baseline). During the return to baseline, stereotypy increased, and engagement and productivity decreased. Given the similar outcomes across DRA and DRO conditions, they were each evaluated individually relative to baseline in subsequent phases. DRA contingencies decreased stereotypy ($M = 16.0\%$ of session relative to 66.2\% in baseline), increased task engagement ($M = 78.5\%$ of session relative to 7.4\% in baseline), and increased productivity ($M = 4.9$ ppm relative to 0.4 ppm in baseline). DRO contingencies also decreased stereotypy ($M = 10.8\%$ of session relative to 56.0\% in baseline), increased task engagement ($M = 68\%$ of session relative to 0.3\% in baseline), and increased productivity ($M = 3.5$ ppm relative to 0.1 ppm in baseline). Although similar decreases in stereotypy occurred during DRA and DRO phases, DRA resulted in greater increases in engagement and productivity than DRO. For this reason, the DRA versus baseline phase was reimplemented. Stereotypy decreased during DRA ($M = 4.3\%$ of session relative to 77.0\% in baseline), engagement increased ($M = 91.9\%$ of session relative to 0.3\% in baseline), and productivity increased ($M = 5.4$ ppm relative to 0.0 ppm in baseline). Following this phase, tokens were delivered on a fixed ratio (FR) 1 schedule with an FR 20 exchange schedule, and treatment outcomes maintained.

**DISCUSSION**

The relative efficacy of DRO and DRA was compared, without response blocking or interruption, for decreasing automatically reinforced stereotypy, sustaining task engagement, and increasing task productivity for three individuals diagnosed with ASD. DRA contingencies reduced stereotypy to below baseline levels for each participant, whereas...
DRO contingencies reduced stereotypy for only one participant. DRA also increased item engagement and productivity to a greater extent than DRO. Thus, the results indicate that DRA resulted in superior treatment outcomes when compared to DRO.

Figure 4. For Solomon, the top panel depicts motor stereotypy, the second engagement, and the third productivity across baseline (BL), differential reinforcement of alternative behavior (DRA) compared to differential reinforcement of other behavior (DRO), return to BL, DRA compared to BL, DRO compared to BL, then a return to DRA compared to BL. During baseline the only difference between the conditions was the color of the materials.
The differential outcomes between DRA and DRO can most easily be attributed to arranging a direct contingency for an appropriate and potentially incompatible behavior (i.e., task engagement) during DRA. In a brief review Jessel and Ingvarsson (2016) discuss mechanisms that are potentially responsible for increasing the effectiveness of DRO. For example, DRO procedures may increase other behaviors than the target behavior through adventitious reinforcement. If DRO procedures are effective due to adventitious reinforcement, repeated exposures to the contingency may be necessary for other behaviors to be strengthened sufficiently to displace stereotypy. By contrast, the repeated reinforcement of a singular alternative behavior during DRA may result in more rapid acquisition of this behavior and a concomitant decrease in the target behavior. Although more prolonged exposure to DRO contingencies may have reduced stereotypy, the speed of attaining reductions via DRA seems preferable.

The DRA conditions may also have resulted in superior outcomes due to differences in the frequency with which reinforcement was delivered. That is, the rate of reinforcement could have been raised under DRA conditions by responding faster, whereas the maximum reinforcement rate in the DRO conditions was fixed. This increased access to, and time spent consuming reinforcement, may have also accounted for differences in interventions. Retrospective review of videos of DRA and DRO sessions revealed the frequency of reinforcer delivery in each condition (see Table 1). DRA resulted in higher reinforcement rates than DRO for John (3.3 reinforcers per minute [rpm] relative to 1.1 rpm, respectively) and Nick (3.9 rpm relative to 2.5 rpm, respectively). Reinforcement rates were roughly equal for Solomon (4.7 rpm relative to 4.4 rpm during DRA and DRO, respectively); this may be due to Solomon’s brief (7 s) omission interval during DRO. Thus, the differential outcomes obtained across DRA and DRO could have been the result of differences in schedules of reinforcement rather than the contingency. To isolate the role of the reinforcement contingency, researchers could control for differences in reinforcement schedule across DRA and DRO by yoking the DRO omission interval to the DRA mean interreinforcement interval (e.g., if reinforcement is delivered every 20 s during DRA sessions, the abstention interval could be set at 20 s during DRO sessions).

During the DRA conditions, reinforcement delivery occurred only following product

<table>
<thead>
<tr>
<th>Phase</th>
<th>John</th>
<th>Nick</th>
<th>Solomon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RDR (rpm)</td>
<td>Stpy (%)</td>
<td>RDR (rpm)</td>
</tr>
<tr>
<td>Total Sessions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRA</td>
<td>3.3</td>
<td>5.6</td>
<td>4</td>
</tr>
<tr>
<td>DRO</td>
<td>1.1</td>
<td>29.7</td>
<td>2.5</td>
</tr>
<tr>
<td>DRA vs. DRO</td>
<td>2.5</td>
<td>7.8</td>
<td>4.8</td>
</tr>
<tr>
<td>DRO</td>
<td>1.3</td>
<td>12.1</td>
<td>3.6</td>
</tr>
<tr>
<td>DRA vs. BL</td>
<td>3.4</td>
<td>6.8</td>
<td>4.2</td>
</tr>
<tr>
<td>DRA</td>
<td>1.2</td>
<td>14.1</td>
<td>2.2</td>
</tr>
<tr>
<td>DRO</td>
<td>4.5</td>
<td>2.1</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Note. Reinforcer Delivery Rate = RDR; Stereotypy = Stpy; Reinforcers per minute (rpm).
completion without problem behavior. These DRA conditions could be conceptualized as a conjunctive DRA:DRO schedule in that both the absence of stereotypy and the alternative response were required to produce reinforcement. This DRA arrangement is commensurate with typical applications; future researchers could further isolate the role of DRO from DRA by delivering reinforcement for task completion, independent of stereotypy.

Although the DRA conditions included omission criteria, the findings are unique in that DRA yielded effective results without response blocking or response interruption procedures. As noted, there have been relatively few demonstrations of DRA contingencies without response blocking. Another feature of the current study is that we targeted a vocational task, whereas previous evaluations of DRA for treating automatically reinforced problem behavior typically have targeted appropriate leisure item engagement (Favell, McGimsey, & Schell, 1982; Vollmer et al., 1994). Stereotypy is frequently identified as a problem behavior to the extent that it competes with engaging in learning experiences. Targeting engagement with leisure items may not address this practical concern to the same extent that targeting direct engagement in vocational and academic tasks would; but the most appropriate alternative activity is likely best dictated by the context from which the client is being referred. Practitioners may increase the social validity of intervention outcomes by allowing caregivers to nominate desired alternative tasks to include in a treatment analysis.

Despite the positive outcomes associated with DRA, the sessions were relatively brief (i.e., 5 min). Practical application would require similar behavioral changes during lengthier exposure to contingencies and under leaner reinforcement schedules. Future research should examine DRA during extended application to determine the maintenance of these intervention effects.

Finally, clinicians might thoroughly consider when stereotypy does and does not constitute a problem behavior for their clients before adopting DRA as a component of their treatment approach. This decision should be made in consultation with clients’ caregivers and the client themselves, if possible. With the exception of repetitive self-injurious behavior, the complete elimination of noninjurious stereotypy is likely not an appropriate goal, as this would restrict access to reinforcement for individuals who may have relatively few reinforcers. Instead, we encourage clinicians to teach their clients to discriminate when the occurrence of stereotypy is problematic relative to when it is harmless (Tiger, Wierzb, Fisher, & Benitez, 2017).

REFERENCES


Received February 19, 2018
Final acceptance March 8, 2019
Action Editor, Jeffrey Tiger