The immediate and subsequent effects of response interruption and redirection on vocal stereotypy, motor stereotypy and heart rate

Tyla M. Frewing a,b, Amy Tanner c, Andrew C. Bonner d, Sharon Baxter b and Sarah J. Pastrana a,b

aDepartment of Educational and Counselling Psychology and Special Education, University of British Columbia, Vancouver, Canada; bSemiahmoo Behaviour Analysts Inc., Surrey, Canada; cCBI-Monarch House, Burnaby, Canada; dArcus Community Resources, Maple Ridge, Canada

ABSTRACT
A three-component multiple-schedule analysis was used to evaluate the immediate and subsequent effects of response interruption and redirection (RIRD) with stimulus control (SC) on levels of motor and vocal stereotypy and heart rate in a 19-year-old male with autism. After a wristband was established as a discriminative stimulus for implementation of RIRD contingent on engagement in motor or vocal stereotypy, the present treatment evaluation was conducted. During baseline sessions, there were no planned consequences for stereotypy. During treatment sessions, RIRD was implemented contingent on each event of targeted stereotypy. Motor and vocal stereotypes were evaluated using 5-s partial interval recording (PIR) and heart rate was monitored via momentary time sampling. Results showed that RIRD with SC produced immediate reductions in motor and vocal stereotypy to near-zero rates during treatment and did not produce subsequent increases in stereotypy when the intervention was removed. Heart rate remained consistent across all conditions. Following the above analysis, generalization probes were carried out in three additional settings. The utility of including SC procedures when implementing RIRD, considerations in using heart rate as a measure of distress, and avenues for future research is discussed.

ARTICLE HISTORY
Received 1 November 2014
Accepted 11 August 2015

KEYWORDS
Stereotypy; response interruption and redirection; stimulus control; automatic reinforcement; heart rate

Stereotypy has been defined as repetitive and invariant movements that occur in the absence of social consequences (Rapp & Vollmer, 2005), and is one of the key diagnostic features of autism spectrum disorder (ASD; American Psychiatric Association, 2013). Engagement in automatically reinforced stereotypy may block the effects of social reinforcers (Lovas, 2003) and the acquisition of new skills (Epstein, Doke, Sajwaj, Sorrell, & Rimmer, 1974; Koegel & Covert, 1972; Lovas, Newsom, & Hickman, 1987; Morrison & Rosales-Ruiz, 1997). Stereotypy maintained by automatic reinforcement can also be difficult to treat because practitioners often do not have access to the maintaining variables (Lovas et al., 1987; Rapp, 2008; Taylor, Hoch, & Weissman, 2005). However, studies have recently identified a number of potentially effective procedures for reducing stereotypy.

CONTACT Amy Tanner amy_tanner@hotmail.com

© 2015 Norwegian Association for Behavior Analysis
One such intervention, a variation of non-contingent reinforcement that involves providing an individual with continuous or fixed-time access to matched stimuli associated with reductions in stereotypy (i.e., stimuli who produce stimulation similar to that produced by engagement in the target form of stereotypy), has been shown to reduce both motor and vocal stereotypy (Lanovaz et al., 2014; LeBlanc, Patel, & Carr, 2000; Rapp & Vollmer, 2005; Rapp et al., 2013). However, the utility of this approach may be limited in settings where engagement with non-contingent stimuli interferes with ongoing activities. For example, it may be difficult or impossible for an individual to complete academic tasks or to engage in social interactions while they manipulate a textured ball, or a toy that produces lights and sounds when buttons are depressed, or other stimuli associated with low levels of stereotypy. In such situations, it may be beneficial to teach individuals “time and place,” such that they learn to inhibit engagement in stereotypy under specific conditions. Bringing stereotypy under stimulus control (SC) such that a discriminative stimulus (S\textsuperscript{D}) signals implementation of an intervention may be one way to achieve this. In a study evaluating the effectiveness of establishing SC over vocal stereotypy, Rapp, Patel, Ghezzi, O’Flaherty, and Titterington (2009) found that when a red card signaled the implementation of verbal reprimands or response cost procedures contingent on engagement in vocal stereotypy, levels of stereotypy decreased in the presence of the red card. Several additional studies have specifically evaluated the effects of response interruption and redirection (RIRD) with SC procedures.

RIRD involves (a) interrupting engagement in stereotypy and (b) prompting the individual to engage in another response (Ahearn, Clark, MacDonald, & Chung, 2007). A number of studies have evaluated the effectiveness of RIRD and found that it decreases engagement in stereotypy (e.g., Ahearn et al., 2007; Ahrens, Lerman, Kodak, Worsdell & Keegan, 2011; Schumacher & Rapp, 2011).

Recent studies evaluating RIRD with SC procedures have signaled the implementation of RIRD with a variety of stimuli. Stimuli have included, for example, a specific therapist (Ahrens et al., 2011), a colored card (Schumacher & Rapp, 2011), or a colored shirt (Love, Miguel, Fernand, & LaBrie, 2012). One potential drawback to these studies is that the stimuli used to signal implementation of the intervention was one that may not easily be used in other settings (e.g., walking in the community, shopping in the grocery store) or people (e.g., other therapists, parents, etc.). Furthermore, RIRD may itself be difficult to implement in generalized settings because (a) it must be implemented upon each occurrence of stereotypy, and (b) emotional responses and other side effects associated with punishment-based procedures may not be tolerable in these settings. SC procedures may provide at least two benefits in such settings: (1) SC procedures may facilitate an individual’s ability to differentiate when RIRD is in effect if RIRD is used only in specific contexts where stereotypy is most disruptive or problematic; and (2) the presence of the antecedent stimuli alone may come to control behavior, such that RIRD needs to be implemented less frequently when the stimulus is present, making it more useful in a variety of settings.

Behavior-reduction interventions like RIRD are often implemented with populations who cannot readily communicate their wants, needs, and experiences to those around them. When individuals receiving treatment have limited communication repertoires,
measures of social validity are important. Most measures of social validity, however, rely on verbal reports provided by recipients of services or their caregivers. There is a notable gap in the assessment of social validity in populations with limited communication skills, given that verbal reports cannot be collected directly from them. When individuals cannot verbally report their experience of an intervention, it may be possible to use physiological measures such as heart rates as an index of arousal and potential distress. Although we were unable to identify any studies that directly assessed physiological measures during the course of intervention for problem behaviors, one study investigated whether certain forms of stereotypy were associated with indicators of elation, distress, and neutral affect or “composure”, as well as whether certain heart rate patterns were associated with stereotypy occurring in the presence of elated, distressed, or composed affect (Willemsen-Swinkels, Buitelaar, Dekker, & Engeland, 1998). Willemsen-Swinkels et al. (1998) found that low-intensity stereotypies producing repetitive sensory input were most often associated with distress, whereas short duration, high-intensity stereotypies were most often associated with elation. Further, the results showed that stereotypy associated with elated, composed, and distressed affect each had a unique heart rate pattern. For example, stereotypy associated with distress showed a peak heart rate just before the onset of stereotypy, with a decrease immediately after it began.

An additional consideration for interventions that aim to suppress stereotypy (e.g., RIRD) is how engagement in stereotypy is affected following intervention withdrawal. If an intervention produces an increase in stereotypy upon removal, it may be less desirable, even if it effectively reduces stereotypy while in place. The three-component multiple-schedule (Lanovaz, Rapp, & Fletcher, 2010) allows for the evaluation of both immediate (i.e., during) and subsequent (i.e., immediately after) effects of an intervention. At least two studies have found that RIRD produces immediate decreases in stereotypy with no subsequent increases (e.g., Pastrana, Rapp, & Frewing, 2013; Schumacher & Rapp, 2011); however, the literature base is limited with regards to the assessment of subsequent engagement in stereotypy.

The purpose of the present study was twofold: First, to evaluate the immediate and subsequent effects of RIRD with SC procedures on the occurrence of motor and vocal stereotypy in an individual with ASD when an easily portable wristband was used to signal implementation of RIRD. Second, to investigate the immediate and subsequent effects of RIRD on heart rate. Heart rate was used as an index of physiological arousal that may indicate the participant experienced some stress associated with the procedure (Utsey & Hook, 2007).

**Method**

**Participants, setting, and materials**

One 19-year-old male diagnosed with ASD participated in the study. John lived in an intensive applied behavior analysis (ABA) teaching home with 24-h care. John had received intensive ABA treatment since the age of 4, spoke in complete sentences and read at approximately a 2nd grade level. All treatment evaluation sessions took place in a small therapy room in the participant’s ABA teaching home. The room contained a
table, two chairs, a bookshelf containing therapy materials, and a filing cabinet. Materials included a wristband (used as the SC item), a digital wristwatch containing a heart rate monitor, and discrete trial teaching (DTT) one-to-one instructional materials. Generalization sessions took place in three different settings. First, the procedure was implemented while the participant listened to a digital handheld music player in the family room of his home. The procedure was then evaluated during mealtimes at the dining room table in his home. Finally, the procedure was evaluated in a “toy room,” which the participant used during leisure time that contained a variety of preferred activities.

**Dependent variables**

Dependent variables included (1) motor stereotypy, (2) vocal stereotypy, and (3) heart rate. Motor stereotypy was defined as any non-functional motor movements (e.g., rocking in chair, slapping belly, hand flapping, finger flicking, leg or belly rubbing, covering and uncovering ears, etc.). Vocal stereotypy was defined as any non-functional or non-contextual vocalization or sound (e.g., laughing, whispering, humming, moaning, immediate and delayed echolalia, etc.) and excluded any functional responses to vocal antecedents delivered by the experimenter. Heart rate was defined as number of beats per minute (BPM) as measured by a digital wristwatch with a heart rate monitor.

**Response measurement and interobserver agreement**

Sessions were videotaped using a tablet device and later scored from video. The occurrence of vocal and motor stereotypy was recorded using 5-s partial interval recording (PIR). An additional observer collected interobserver agreement (IOA) data for 20% of sessions for the baseline sequences and 40% of sessions for the treatment sequences. Observers were provided with operational definitions of motor and vocal stereotypy and trained to identify examples and non-examples in videos. Observers then scored 10-min sample videos and results were compared to a trained observer. Training continued until observer agreement reached 90% or higher across two 10-min sessions. IOA was calculated interval-by-interval, by dividing the number of intervals with agreements by the number of intervals with agreements plus disagreements and multiplying by 100. The mean IOA across all sessions was 90% (range, 88–92%). Heart rate was monitored using momentary time sampling. Once every 2.5 min, the participants’ heart rate was recorded by pressing a button on the monitor. Heart rate data was scored via paper and pencil by the experimenter, and no interobserver data were collected for heart rate. During generalization probes, data were collected using 5-s PIR.

**Experimental design**

A multielement design with an embedded three-component multiple-schedule (Lanovaz et al., 2010) was used to evaluate the immediate and subsequent effects of RIRD and SC procedures on motor stereotypy, vocal stereotypy, and heart rate. Generalization probes were conducted to evaluate the effectiveness of the RIRD and
SC procedures in three additional settings. A 5-min baseline probe and a 5-min treatment probe were conducted in each setting.

**Procedures**

**Pre-teaching**
Prior to the treatment analysis, the RIRD with SC intervention was implemented in initially short (3-min) intervals. Intervals were then gradually and systematically increased by short durations (e.g., 15–30 s), each time the participant met mastery criterion (two times consecutively at 90–100% of intervals with an absence of stereotypy and compliance with RIRD) to a terminal interval of 20-min duration.

**General procedures**
Each sequence contained three consecutive 10-min components. The experimenter implemented DTT maintenance programs (i.e., requiring mastered skills only) during all sessions.

**Baseline sequence**
Procedures remained the same throughout all three 10-min components. No planned consequences were provided for motor and vocal stereotypy.

**Treatment sequence**
During the treatment sequence, baseline procedures were implemented in the 1st and 3rd component, and no planned consequences were delivered for stereotypy. During the second component, the RIRD with SC procedure was implemented. During the intervention, the participant wore the SC wristband on the right wrist, and a vocal discriminative stimulus was delivered: “Hands, body, noises” (an abbreviation for “Calm hands, calm body, no noises.”). When the SC wristband was on, RIRD was implemented for each event of targeted stereotypy.

**RIRD procedures**
Contingent on each event of vocal stereotypy, the participant was directed to engage in an incompatible vocalization: He was directed to count from 100 to 200. The timer for the current interval was paused while the RIRD intervention was implemented as this procedure took up to 30 s to complete. Contingent on each event of motor stereotypy, the participant was directed to engage in incompatible motor movements (e.g., put hands in pockets). The interval was not paused when RIRD was implemented for motor stereotypy, as this procedure typically took under 2 s to be completed. If, at any time, the participant did not comply with the instruction to engage in an incompatible behavior, the instruction was then re-presented. If the participant did not comply with the reissued instruction, a calm-down procedure would have been implemented as per his behavior management plan. During the course of this study, the participant always complied with RIRD instructions the first time they were delivered.
**Generalization probes**

Procedures for generalization probes were identical to those during the RIRD with SC component of the treatment sequence, except that they were conducted in a generalized setting. In the first two settings (music device and mealtimes), the procedure was initially introduced with short (3 min) intervals that were systematically increased to 15 min before generalization probes were conducted. For the final setting (toy room), the baseline and generalization probes were conducted with no additional teaching.

**Data analysis**

The three-component multiple-schedule analysis used in the present study produces two possible avenues of data analysis: a between-sequence analysis and a within-sequence analysis (Lanovaz et al., 2010). The between-sequence analysis provides data regarding the immediate treatment effect(s) by comparing responding across each component of the baseline and treatment sequences. If a treatment effect occurs, the treatment sequence should mirror the baseline sequence in the first and third components, but should be differentiated in the second (i.e., treatment) component.

In the within-sequence analysis, the first and third components for each sequence are compared to evaluate the subsequent effects of the treatment. At least three patterns of responding can be observed. First, levels of stereotypy may be consistently higher in the third component (i.e., post-treatment) than in the first component (i.e., pre-treatment) of the treatment sequence. This pattern of responding would indicate that implementation of the treatment produces an establishing operation for the reinforcers produced by engagement in stereotypy. Second, stereotypy may be consistently lower in the third component than the first component. This pattern indicates the treatment may create an abolishing operation for the reinforcers produced by engagement in stereotypy. Finally, levels of stereotypy in the first and third components of the treatment sequence may be similar indicating that the treatment (a) does not produce an abolishing or establishing operation and (b) does not have any lasting effects on the stereotypy once the treatment is removed.

**Results**

When examined separately, the results for motor and vocal stereotypy showed the same pattern of responding as was observed when the results were combined. Thus, given that presenting the results individually provides little additional information, for the remainder of this report, “stereotypy” will refer to the combined results for motor and vocal stereotypy. Figure 1 shows the results for stereotypy. In the between-sequence analysis (top panel), stereotypy occurred at high and stable levels across all three components of the baseline sequence. Similarly, levels of stereotypy were high and stable in the first ($M = 79\%$) and third ($M = 79\%$) components. During the second component of the treatment sequence in which RIRD with SC was implemented, stereotypy reduced to near-zero rates ($M = 2\%$) in the treatment sequence. These results indicate that RIRD with SC effectively reduced stereotypy.

In the within-sequence analysis (bottom panel) for stereotypy, results showed that stereotypy was marginally lower in the third component than in the first for 2/5
treatment sequences and marginally higher in the third component than the first for 3/5 sequences. Further, in all treatment sequences levels of stereotypy were the lowest in the second component when RIRD with SC was implemented. Together, the results of the between- and within-sequence analysis for stereotypy indicate that RIRD with SC produced an immediate decrease in stereotypy and did not produce any subsequent increases in stereotypy.

Figure 2 depicts the results for heart rate. In the between-sequence analysis (top panel) heart rate remained stable within each sequence and was marginally lower during treatment sequences ($M = 81.08$) than during baseline sequences ($M = 93.97$). Given that this difference was consistent across all three components and thus, heart rate was slightly lower in the treatment sequence regardless of whether it was a baseline session (first and third components) or treatment session (second component); these results indicate that implementation of RIRD with SC did not affect heart rate. In the
within-sequence analysis (bottom panel) of heart rate, the heart rate was marginally higher in the third component than the first component with the mean difference between the first and third components being 5.8 BPM.

Following the treatment evaluation described above, generalization probes were conducted to evaluate the RIRD with SC procedure in three additional settings. Table 1 depicts these results. During mealtime, stereotypy occurred in 43% of intervals in the baseline probe and 0% of intervals during the treatment probe. During the music device probe, stereotypy occurred during 95% of intervals in baseline and decreased to 8% during the treatment probe. Finally, in the toy-room probe, stereotypy occurred during 48% of intervals during baseline, and decreased to 0% of intervals when the intervention was implemented. Together, these results indicate the RIRD with SC intervention was effective in reducing stereotypy in three additional environments.

Figure 2. The between-sequence analysis (top panel) shows heart rate (BPM) during the first, second, and third component of the baseline and treatment sequences. The within-sequence analysis (bottom panel) shows heart rate (BPM) within each sequence. Data were omitted from sessions 1 and 2 of treatment due to mechanical error with heart rate instrument. Each sequence contains three successive components: component 1 (lightest bar, first bar in each sequence), component 2 (black bar, second bar in each sequence), and component 3 (dark gray bar, third bar in each sequence).
Discussion

The present study found RIRD with SC produced immediate decreases in stereotypy and no subsequent increases or decreases when the intervention was removed. The current study replicates previous findings that RIRD is an effective intervention for stereotypy (e.g., Ahearn et al., 2007; Lerman & Iwata, 1996; Pastrana et al., 2013; Schumacher & Rapp, 2011). Additionally, results are similar to Pastrana et al. (2013) and Schumacher and Rapp (2011) in that RIRD did not produce increases in stereotypy after it was removed. A unique contribution of this study was the utilization of a heart rate monitor to evaluate both immediate and subsequent effects of RIRD and SC on heart rate. In the present study, the participant’s heart rate did not increase when RIRD with SC was implemented or immediately after it was removed. These results provide preliminary evidence that it is unlikely the participant experienced distress as a result of the intervention. Heart rate results should be interpreted with caution, however, for at least three reasons: First, the tool used to measure heart rate was rudimentary and given the small differences observed, a more accurate tool may detect greater variability between conditions. Second, in the present evaluation, heart rate was only recorded once every 2.5 min. Given that at least one previous study (Willemsen-Swinkels et al., 1998) found that heart rate peaked just before the onset of stereotypy in 85% of stereotypic behaviors associated with displays of distress (indicated by crying, screaming, negative vocalizations, etc.), it would be valuable to utilize a heart rate monitor that provided more frequent evaluations (e.g., once every 5 s). Third, the motor stereotypy emitted by the participant involved repetitive large body movements such as rocking. These movements may produce an elevated heart rate during baseline, possibly obscuring results.

There are some limitations to the present study. First, a functional analysis was not conducted prior to starting treatment and therefore the function of the stereotypy was not conclusively determined. However, the persistence of stereotypy in the absence of any planned social consequences during the baseline sequence provides some evidence that it is automatically maintained (Vollmer, Marcus, Ringdahl, & Roane, 1995). Second, treatment effects were not long lasting. Thus, once the intervention was removed, levels of stereotypy quickly returned to baseline levels. Third, treatment integrity data were not collected. Finally, appropriate vocalizations were not measured in this study, and it is not possible to evaluate whether RIRD with SC increased, decreased, or had no effect on appropriate vocalizations. Pastrana et al. (2013) investigated whether RIRD intervention for one target form of stereotypy produced changes in levels of another untargeted behavior. Future studies might use a similar method to investigate the effects of RIRD for stereotypy on appropriate vocalizations.

Table 1. Generalization probes.

<table>
<thead>
<tr>
<th>Setting</th>
<th>Baseline</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mealtime</td>
<td>43.3%</td>
<td>0%</td>
</tr>
<tr>
<td>Music Device</td>
<td>95.0%</td>
<td>8.3%</td>
</tr>
<tr>
<td>Toy Room</td>
<td>48.3%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Note: Percentage of time spent engaged in stereotypy during 5-min probes. Pre-teaching occurred in Mealtime and Music Device settings. No pre-teaching occurred in toy room, and data were collected the first time treatment was implemented in this setting.
The present study provides a starting point for the possible inclusion of measures of physiological arousal when evaluating treatment effectiveness. Physiological measures may have the greatest potential for use in populations who cannot effectively communicate distress or discomfort to those around them. This may be of particular value when practitioners, caregivers, or other stakeholders are concerned that an individual may experience distress as a result of an intervention, or, when an intervention utilizing punishment-based procedures is implemented. Alongside other methods (e.g., evaluating preference via a concurrent operants preparation), measuring physiological indexes of arousal may be one way to directly assess an individual’s experience of an intervention. Future research should investigate heart rate results when a more sophisticated monitor is used with more frequent heart rate measurements.

The present study also evaluated the use of a SC item that was easily generalized to multiple settings. This type of stimulus (i.e., a small wristband), though discrete and easily portable may not be sufficiently salient in some cases and stimulus fading procedures may be required. Future research should consider the inclusion of a measure for the number of RIRD implementations required when RIRD is implemented with SC procedures, to evaluate whether the inclusion of a SC item reduces the number of required RIRD implementations. Further, researchers might evaluate whether stereotypy remains suppressed when the schedule of RIRD implementation is thinned (i.e., if it is implemented less consistently) in the presence of a SC item.

Disclosure statement

No potential conflict of interest was reported by the authors.

References


