12-2020

A Component Analysis of the Effects of Response Interruption and Redirection on Vocal Stereotypy in an Adult with Autism Spectrum Disorder

Andrew H. Chen
St. Cloud State University, je4268wj@go.minnstate.edu

A Component Analysis of the Effects of Response Interruption and Redirection on Vocal Stereotypy in an Adult with Autism Spectrum Disorder

by

Andrew H. Chen

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State university

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in Applied Behaviour Analysis

December, 2021

Thesis Committee:
Michele Traub, Chairperson
Benjamin Witts,
Kimberly Schulze

Abstract

There has been much research into evaluating the effectiveness of response interruption and redirection (RIRD) in the reduction of vocal stereotypy in children and adolescents diagnosed with autism spectrum disorder (ASD). Research has indicated that RIRD often results in the reduction in level of vocal stereotypy in this population. However, only one previous study has evaluated the efficacy of RIRD on vocal stereotypy for participants older than 18 years old. Furthermore, though some studies point to punishment as the mechanism by which RIRD produces its effects (Ahearn et al., 2007; Aherns et al., 2011), it is still described as a redirection procedure with unclear contingencies (Cassella et al., 2011). This study used the uninterrupted data collection procedures described by Carroll and Kodak (2014) and Wunderlich and Vollmer (2015) which have been shown to provide a more accurate analysis compared to the interrupted technique. Additionally, this study replicated and expanded upon Wunderlich and Vollmer (2015) by introducing a component analysis of the effects of RIRD on an adult participant. The results showed that motor RIRD was effective in reducing the vocal stereotypy, that random talking may be an establishing operation for vocal stereotypy, and that levels of appropriate vocalizations, while initially supressed for 12 sessions, did not change meaningfully throughout the study.

Acknowledgements

I would like to extend my most sincere thank you to my advisor and committee chair, Dr. Michele Traub, whose input and guidance contributed significantly to the project. Thank you for being an amazing advisor and guiding me through this project every step of the way, without your guidance this project would not have been possible.

A special thanks goes out to my supervisor Andy Harrison and colleague Terry Rogers who let me bounce ideas off of and helped me with data collection.

Lastly, I would like to thank my parents, Ku Shu-Chin and Chen Ko-Ju whose love, patience and encouragement made this journey possible. I love you.

Table of Contents

Chapter I: Introduction

The assessment and treatment of stereotypic behaviour, or stereotypy, has been a focus of behaviour analytic research for many years (Wunderlich & Vollmer, 2015). While there is no consensus on the exact definition of stereotypy (Rapp & Vollmer, 2005), it can be generally defined as repeated motor or vocal behaviour that serves no social function (Rapp et el., 2004). While stereotypy can be observed to occur in typically developing children and adults, it is a core diagnostic criterion for individuals diagnosed with an Autism Spectrum Disorder (ASD) (American Psychiatric Association, 2013).

Topographies of stereotypy are highly idiosyncratic across individuals, but common examples of motor stereotypy include toe walking, hand flapping, body rocking, swaying, object manipulation, and even self-injurious behaviour. For vocal stereotypy, common examples include humming, singing, guttural noises and the reciting of dialogue. Stereotypy may occur frequently, involve large motor movements that can be dangerous, or may involve loud disruptive singing lasting many seconds in duration (Lydon et al., 2013). Stereotypy can affect skill acquisition and impede opportunities to learn and interact appropriately with the environment. Furthermore, vocal stereotypy directly affects communication and interactions with peers and can be socially stigmatizing (Ahrens et al., 2011).

It is important to note the classification of stereotypy as an operant rather than a reflex, and that functional analysis of stereotypy most often demonstrates it to be maintained via automatic reinforcement (Rapp & Vollmer, 2005; Taylor et al., 2005). Stereotypy is a voluntary behaviour that is maintained by the consequences that follow it. Presumably, stereotypic behaviours are maintained by the perceptual or sensory consequences they produce (Kennedy et

al., 2000). These non-social reinforcement contingencies are challenging to treat because the reinforcing consequences are often inaccessible (Vollmer, 1994) and, therefore, cannot be manipulated.

Ahearn et al. (2007) evaluated the effects of a response blocking procedure, called response interruption and redirection (RIRD), in the treatment of vocal stereotypy in four participants diagnosed with ASD. RIRD is an intervention that delivers demands contingent on stereotypy; during treatment sessions, vocal demands were presented contingent on the occurrence of vocal stereotypy and were continuously provided until the participants complied with three consecutive responses without engaging in vocal stereotypy. Praise was given at the end of a successful RIRD sequence for engaging in appropriate vocal responses. Results showed that RIRD significantly lowered the levels of vocal stereotypy for all four participants compared to baseline. The authors suggested two methods by which RIRD reduced vocal stereotypy: either by reducing contact with sensory consequences maintaining the behaviour or by directly punishing the stereotypic responses.

Ahrens et al. (2011) conducted a series of experiments to determine the mechanisms of the reductive effects of RIRD suggested by Ahearn et al. (2007) by comparing the effects of both vocal and motor RIRD contingent on the occurrence of vocal stereotypy. In the first experiment, the authors found similar effectiveness between motor RIRD and vocal RIRD on the level of vocal stereotypy for two boys diagnosed with ASD. In the second experiment, the authors again found similar effectiveness for both motor and vocal RIRD on vocal stereotypy in addition to motor stereotypy for two additional participants. These results suggested that RIRD was effective regardless of the topography of the demand. Notably, Ahrens et al. (2011)

terminated vocal RIRD after the third demand, regardless of compliance; despite this change from the Ahearn et al. (2007) procedure, vocal RIRD was equally effective for noncompliant participants.

Cassella et al. (2011) systematically replicated and extended Ahearn et al. (2007) with two boys diagnosed with ASD. Cassella et al. used an ABAB reversal experimental design to evaluate the effects of motor RIRD on the level of vocal stereotypy. Cassella et al. found motor RIRD to be effective at reducing the percentage of intervals with vocal stereotypy, which was consistent with the findings of Ahrens et al. (2011).

Shawler and Miguel (2015) used a multiple treatment reversal design to compare the effectiveness of both vocal RIRD and motor RIRD on vocal stereotypy and appropriate vocalizations. They found that for four of their five participants, both demand topographies were similarly effective in reducing vocal stereotypy. These results are consistent with prior research on motor RIRD and suggest that the effectiveness of RIRD is likely not due to the topographical incompatibility of the demand with the stereotypy (i.e., vocal RIRD for vocal stereotypy). This suggests that RIRD is a punishment procedure (the other potential mechanism proposed by Ahearn et al., 2007).

In the context of RIRD research, appropriate responses are usually defined as contextually appropriate vocalizations such as requests, tacts or interverbals. Tracking the frequency of appropriate vocalizations during RIRD implementation sessions is important both because vocal stereotypy may directly compete with appropriate vocalizations and because, if RIRD does in fact function as a punishment procedure, appropriate vocalizations may be punished along with the stereotypic responses. In either event, if reduction in vocal stereotypy

leads to increased appropriate vocalizations, then the individual may not require specific programming or further intervention beyond RIRD to see increases in appropriate vocalizations.

However, data on the effect of RIRD on appropriate responses have shown mixed results. While Ahearn et al. (2007) found that vocal RIRD increased the rate of appropriate vocalizations in three of their four participants, Ahrens et al. (2011) reported mixed results for two of their four participants (e.g., unclear or low increases), Cassella et al. (2011) were unable to increase appropriate vocalizations with motor RIRD, and Shawler and Miguel (2015) found that both motor and vocal RIRD were equally effective in increasing appropriate vocalizations. Thus, more research is required to determine the effects of RIRD on appropriate vocalizations.

Though the findings of Ahearn et al. (2007) were significant enough to warrant increased research and focus, their data collection procedures may have included measurement artifacts that overestimated the efficacy of RIRD. The procedure used by Ahearn et al. (2007) involved time stoppage during intervention which, though consistent with previous overcorrection-based procedures (Epstein et al., 1974, as discussed in Wunderlich & Vollmer, 2015), may have underestimated the frequency and duration of vocal stereotypy.

To this end, Carroll and Kodak (2014) examined the effects of interrupted and uninterrupted measurement strategies on the effectiveness of RIRD in two boys, 5-year-old and 8-year-old. They presented motor demands for vocal stereotypy. For both participants, they found that when the uninterrupted measurement method was used (that is, when vocal stereotypy during RIRD implementation was counted) there was a modest reduction in vocal stereotypy during RIRD sessions. However, when the interrupted measurement method was used a significant reduction in vocal stereotypy was observed. The findings of Carroll and Kodak

(2014) suggested that interrupted measurement procedures may overestimate the effectiveness of RIRD as vocal stereotypy that occurred during the RIRD demand sequence is not recorded. For both participants, when vocal stereotypy during RIRD was measured using the uninterrupted method, time outside of RIRD implementation was low and time in RIRD was high, suggesting that much of the session was spent in the RIRD demand sequence.

Similarly, Wunderlich and Vollmer (2015) examined the difference in data collection methods on the effectiveness of RIRD in seven participants. They found mixed levels of reduction when examining the data of the entirety of the session, suggesting an overestimation of the efficacy of RIRD by previous studies (Ahearn et al., 2011; Ahrens et al., 2007; Cassella et al., 2011; Shawler & Miguel, 2015) in line with the findings of Carroll and Kodak (2014).

Lastly, a gap in the literature exists on the effectiveness of RIRD on adult participants. While there has been increased research on the effects of RIRD on stereotypy in children and adolescents, few have explored the effects of RIRD on adult populations. Children with ASD who engage in vocal stereotypy typically continue to display stereotypy throughout adolescence and adulthood (Cunningham & Schreibman, 2008). This suggests that stereotypy can continue to impact leisure, vocational and volunteer opportunities. In these settings the stigma surrounding the behaviour may be magnified due to the perceived inappropriateness of the behaviour due to age.

In a review of the 15 studies that target stereotypy identified by Lydon et al. (2013) and Wong et al. (2014), no RIRD investigations targeting vocal stereotypy were found that included participants that were older than 11-years-old (Wells et al., 2016). Only recently did Wunderlich and Vollmer (2015) include a participant over 18-years-old in their analysis of the effects of

RIRD. It is important to note that, for the 20-year-old adult participant, Wunderlich and Vollmer (2015) found that both motor and vocal RIRD were similarly effective at reducing vocal stereotypy.

It is clear that more research on the effects of compliance of RIRD sequence on vocal stereotypy (Ahearn et al., 2011) and the effect of RIRD on vocal stereotypy and appropriate vocalizations in adult populations is needed (Wunderlich & Vollmer, 2015). Thus, the purpose of this study is to extend the literature on RIRD to adult populations by applying the procedure to vocal stereotypy in an adult participant. The current study also aims to examine the impact of demand compliance on RIRD effectiveness and the effect of RIRD on appropriate vocalizations.

Chapter II: Method

Subject and Setting

One adult diagnosed with ASD participated in this study. Eric (pseudonym) was a 30-year-old man who communicated using a few words or simple sentences. His vocal stereotypy was in the form of loud and sometimes repetitive guttural utterances, humming and sustained yells.

Response Measurement and Interobserver Agreement

Vocal Stereotypy

Vocal stereotypy was defined as any occurrence of noncontexual vocalizations including, but not limited to, repetitive noises, nonspeech sounds (i.e, 'waaaaaahhh', 'triptraptriptrap'), guttural noises, noncontextual laughter, or singing. As the duration of each occurrence of vocal stereotypy varied across responses, vocal stereotypy was recorded as a continuous duration measure with an onset-offset criterion of 1 second. The percentage of each session in which a subject engages in vocal stereotypy was calculated by dividing the number of seconds engaged in vocal stereotypy by the number of seconds in the session then multiplied by one hundred and converting the result to a percentage.

Appropriate Vocalizations

Appropriate vocalizations were defined as contextually appropriate verbal responses, such as mands for tangibles, attention, or edibles of which most vocalizations for the participant fall under this category; tacts; and appropriate vocal requests for information on staff and scheduling. As appropriate vocalizations usually occurred as discrete instances of behaviour, a frequency tally was used.

Data were collected, analyzed, and displayed according to the method described by Carroll and Kodak (2014). Data were presented based on the total percentage of the session in which the subject engaged in vocal stereotypy, including vocal stereotypy that occurred during RIRD demand sequences. This data collection procedure was more precise than the procedure used by Ahearn et al. (2007) in demonstrating the effectiveness of RIRD on vocal stereotypy.

To assess interobserver agreement on the occurrence or non-occurrence of the target behaviour in the first baseline phase, two observers, the first author and a secondary observer (the clinical supervisor of the teaching home), independently scored data for at least 33% of sessions. Due to COVID-19 precautions, all sessions after the first baseline phase were recorded by an Ipad 2 placed on the table facing the participant and the primary experimenter. The secondary observer scored exclusively from video after the first baseline phase.

Data were collected using a data collection application "Countee" on each observers' mobile phones. Data collected by the two observers were collected and scored for agreement and disagreement. An agreement was recorded if both observers recorded an occurrence or non-occurrence, a disagreement was recorded otherwise. The number of 1-s intervals in which the observers agreed on either the occurrence or non-occurrence of the behaviour was divided by the total number of seconds in the session and the result was converted to a percentage.

Reliability data were collected during 47% of all baseline sessions, and 36% of all treatment sessions. Mean agreement for all sessions was 93% (range, 82% to 99%), mean agreement for baseline sessions was 94% (range 87% to 97%), mean agreement for RIRD sessions was 98% (range, 96% to 99%), and mean agreement during random talking session was 85% (range, 82% to 87%). Mean agreement for appropriate vocalization was 100%.

A functional analysis was not conducted prior to this experiment as there is enough agreement in the literature to suggest that stereotypy is predominately reinforced by automatic reinforcement. Furthermore, the baseline of the experiment (Phase A) was set up in a way that is identical to an alone condition in a functional analysis. If vocal stereotypy occurs during the baseline, then it confirms that the behaviour occurs in absence of social contingencies, obviating the need for a full functional analysis.

Chapter III: Procedure

Experimental Design

Component analysis of the RIRD procedure was evaluated using an ABABC withdrawal design with an embedded alternating treatments design within the treatment condition and an extended RIRD-only component. During the treatment phase, an RIRD session was followed by a random talking session. To start all sessions, the participant was asked to stand up from a seated position, followed by the experimenter taking the chair, sitting down, and starting the timer in view of the Ipad camera which signalled the start of the session for data collection purposes.

Baseline

Baseline sessions lasted 5 minutes, during which the participant was free to exhibit any behaviour, all of which were ignored by the experimenter. Subsequent baseline sessions in the second baseline phase were conducted in an identical manner to the first.

Treatment

Response Interruption and Redirection

For treatment sessions, both RIRD sessions and random talking sessions lasted 5 minutes and all behaviours aside from appropriate vocalizations and vocal stereotypy were ignored. If the participant engaged in an appropriate vocalization, such as a mand, the therapist responded positively but withheld the item or activity (e.g., “Nice asking, but you need to wait.”). During RIRD sessions when the participant engaged in vocal stereotypy, the experimenter issued three consecutive motor demands (e.g., “Touch your nose,” “Touch your knees,” “Touch your ears”.). If the participant responded correctly and without emitting further vocal stereotypy during the

demand sequence, then the demand sequence was terminated. If the participant emitted vocal stereotypy during the sequence, three additional motor demands were issued; that is, the termination of the motor demand sequence was contingent on the completion of three demands without occurrence of vocal stereotypy. Subsequent RIRD sessions in the RIRD-only phase were identical to the RIRD sessions during the treatment phase.

Random Talking

During random talking sessions, contingent on vocal stereotypy, the experimenter would talk in full sentences on any number of random topics for a predetermined time (see Appendix for a list of phrases). If vocal stereotypy occurred during this random talking sequence, an additional average duration would be added. The duration of a random talking sequence as well as the average duration of random talking was yoked to the previous RIRD session. To do this, the duration of the RIRD demand sequence in the previous RIRD session was scored. Using a python script (Bangladesh Premier League, 2016, see Appendix A, Figure 1), a list of durations of random talking was generated from the average duration of RIRD demand sequence implementations in the previous session using the range as well as the number of implementations. For example, if the average duration of the previous RIRD session was five seconds, the range of the demand sequence was 4s-16s, and there were 10 implementations then the script would generate a randomized list of 10 durations that is both range bound and average bound. The termination of random talking was contingent on the absence of vocal stereotypy for the duration of the random talking sequence.

Chapter IV: Results and Discussion

During baseline, the percentage of the session with vocal stereotypy was in the 10.7% to 37.0% range with an average of 22.4%. During RIRD sessions, the percentage of the session with vocal stereotypy had a range of 1.7% to 16% with an average of 5.4%. In the random talking sessions vocal stereotypy occurred in the range of 6.0% to 46.0% with a mean of 23.9% (See Appendix A, Figure 2).

While the number of appropriate vocalizations did not change significantly through each phase (See Appendix A, Figure 3), the level was briefly suppressed during the first treatment phase for 12 sessions before returning to the usual levels.

The percentage of vocal stereotypy in RIRD sessions was notably lower than in baseline, which is consistent with findings from previous studies (Ahearn et al., 2007; Aherns et al., 2011; Casella et al., 2011; Shawler & Miguel, 2015) showing that RIRD implementation decreased levels of vocal stereotypy. Additionally, the use of motor RIRD to decrease vocal stereotypy to avoid the use of incompatible responses in an adult participant is consistent with the findings of Wunderlich and Vollmer (2015).

The purpose of the presentation of random talking contingent upon vocal stereotypy was to determine whether contingent neutral-sound presentation by itself (i.e., interruption) could reduce the participant's vocal stereotypy, or if the redirection component is specifically needed. An interruption procedure may be effective because a competing stimulus takes the place of an automatic stimulus and directly disrupts the behaviour. If this is true, then a contingent neutral sound stimulus would be effective in lowering levels of vocal stereotypy. However, in this condition, vocal stereotypy during random talking sessions remained at a variable but elevated

level, similar to baseline before and after withdrawal, suggesting that neutral vocal verbal behaviour did not 'block' nor 'redirect' the stimulus for automatic reinforcement. This suggests that the demand (and thus the effectiveness of RIRD) may be contingent on the aversiveness or response effort of the demand aspect of RIRD, rather than the vocal aspect of the RIRD.

During the first treatment phase, it took approximately three of each RIRD and random talking sessions before levels of vocal stereotypy diverged and became separated. Furthermore, in the first three sessions of the random talking condition, vocal stereotypy levels fell to levels similar to the RIRD condition before returning to baseline.This result was not replicated in the reimplementation of treatment after the withdrawal, as the data show an immediate separation of vocal stereotypy levels when comparing RIRD sessions to random talking sessions. This suggests that for those first three sessions, random talking may have functioned as a punisher for vocal stereotypy. One possible explanation for this is that in these initial random talking sessions, the participant actively listened for the presentation of demands and when it became clear that the talking was 'random', vocal stereotypy levels increased to baseline levels.

Appropriate vocalization was initially suppressed for 13 sessions, and it should be noted that appropriate vocalizations were supressed for approximately twice as many sessions compared to vocal stereotypy (Figure 3). An explanation for this phenomenon is that random talking may have briefly generalized and became a conditioned punisher which lead to a decrease in vocal verbal behaviour. The suppression of appropriate vocalizations may not come as a surprise, since most of the participant's appropriate vocalizations came in the form of a mand, which was met with social praise, but access to the item was delayed until the end of the session. It is possible that the initial implementation of treatment caused all vocal-verbal

behaviours from the experimenter to become a conditioned punisher, but recovery of the vocalizations that were automatically reinforced was quicker. This is presumably because RIRD is a punishment technique and not a 'blocking' nor 'redirection' technique that leads to sensory extinction. This means that pathways and contingencies that accessed automatic reinforcers were not severed (read: not severable via RIRD), while vocalizations such as appropriate vocalizations that did not access powerful reinforcers required more sessions before the levels returned to baseline.

Another interesting result is that average vocal stereotypy was 1.5% higher (6.7% more) in the random talking phase than in baseline, with a higher maximum range (37% compared to 46%). This suggests that random talking may have properties of an establishing operation and functioned as events that increase the reinforcing effectiveness of vocal stereotypy. For example, if the sensory reinforcer for vocal stereotypy is sound, and an increase in the amount of sound in an environment was contingent on engaging in vocal stereotypy (contingent random talking), then it follows that random talking would reinforce (i.e., increase) vocal stereotypy.

Lastly, vocal stereotypy levels remained low in the extended RIRD only phase, which suggests that the suppression of vocal stereotypy was not due to any unforeseen concurrent schedules of punishment during the treatment phase.

Limitations and Future Research

One limitation is that the number of subjects in the current study is too few to generalize the results to a population of adults with vocal stereotypy. The effects of the demand component of RIRD on vocal stereotypy in the single participant of this study, while shown as effective in

the reversal, was not replicated in other individuals. Future research should seek to replicate the findings of this study with other adult participants.

Another limitation was that the python script did not generate values that were hard bound to the average duration. For example, if the average RIRD implementation duration was five seconds in the previous session, the script may generate a set of values that averaged 5.7 seconds. Likely the code was insufficiently developed to generate a true random sequence of numbers that is both range bound and average bound, as being both range bound and average bound necessarily constraints the variables needed for a random generation of values. Nevertheless, each set of values was checked after it was generated to ensure that the average did not deviate more than one second more or less than the average RIRD duration from the previous session.

Future research might evaluate other dimensions of vocal stereotypy was decreased. While the percentage of vocal stereotypy during RIRD sessions has decreased, data were not collected on whether the loudness of vocal stereotypy has decreased. It may be more pragmatic to parents and educators of adult participants to present RIRD as a technology to decrease loud, disruptive vocal stereotypy for a short duration of time. For an adult with ASD, a reduced or suppressed level of vocal stereotypy can be useful, even if it is a short duration, if it allows the individual to attend social activities with family or friends such as going to the movies, or eating at a restaurant without being stigmatized.

Another avenue of research might evaluate the use of more covert motor demands and its effect on levels of vocal stereotypy. If the aim of the therapy is to decrease possible stigmatization in public spaces then perhaps one way of doing so is to reduce the range motor

movements necessary. “Close your eyes,” “Wiggle your nose,” “Smile,” “Open your eyes” may be more appropriate than “Touch your toes,” “Touch your head,” and “Clap your hands”. If the response effort of these more covert motor demands is similarly effective in reducing levels of vocal stereotypy then it may be more socially acceptable and palatable to parents and educators to use RIRD in public spaces.

Lastly, future research might expand on this study by presenting RIRD demands in a different modality to investigate the effects of other components in the RIRD demand sequence. For example, motor demands may be presented via flash cards to eliminate the need for vocal instructions from the instructor. This may increase the flexibility of RIRD as a technology that can be used in situations where discreetness is valuable.

References

Ahearn, W. H., Clark, K. M., MacDonald, R. P. F., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behaviour Analysis, 40,* 263-275. doi: 10.1901/jaba.2007.30-06

Ahrens, E. N., Lerman, D. C., Kodak, T., Worsdell, A. S., & Keegan, C. (2011). Further evaluation of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behaviour Analysis, 44,* 95-108. doi: 10.1901/jaba.2011.44-95

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Bangladesh Premier League [BPL]. (2016, September 11). How to generate numbers in range with specific average with Python? [Online forum post]. *Stack Overflow*. https://stackoverflow.com/questions/39435481/how-to-generate-numbers-in-range-with-specific-average-with-python

Carroll, R. A., & Kodak, T. (2014). An evaluation of interrupted and uninterrupted measurement of vocal stereotypy on perceived treatment outcomes. *Journal of Applied Behaviour Analysis, 47,* 1-13. doi: 10.1002/jaba.118

Cassella, M. D., Sidener, T. A., Sidener, D. W., & Progar, P. R. (2011). Response interruption and redirection for vocal stereotypy in children with autism: a systematic replication. *Journal of Applied Behaviour Analysis, 44,* 169-173. doi: 10.1901/jaba.2011.44-169

Cunningham, A. B., & Schreibman, L. (2008). Stereotypy in autism: The importance of function. *Research in Autism Spectrum Disorders, 2,* 469-479. doi: 10.1016/j.rasd2007.09.006

Epstein, L. H., Doke, L. A., Sajwaj, T. E., Sorrell, S., & Rimmer, B. (1974). Generality and side effects of overcorrection. *Journal of Applied Behaviour Analysis, 7,* 385-390. doi: 10.1901/jaba.1974.7-385

Kennedy, C. H., Meyer, K. A., Knowles, T., & Shukla, S. (2000). Analyzing the multiple functions of stereotypical behaviour for students with autism: Implications for assessment and treatment. *Journal of Applied Behaviour Analysis, 33,* 559-571.

Lydon, S., Healy, O., O'Reilly, M., & McCoy, A. (2013). A systematic review and evaluation of response redirection as a treatment for challenging behaviour in individuals with developmental disabilities. *Research in Developmental Disabilities, 34,* 3148-3158. doi:10.1016/j.ridd.2013.06.010

Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: A review of behavioural assessment and treatment. *Research in Developmental Disabilities, 26,* 527-547. doi: 10.1016/j.ridd.2004.11.005

Rapp, J. T., Vollmer, T. R., St. Peter, C., Dozier, C. L. & Cotnoir, N. M. (2004). Analysis of response allocation in individuals with multiple forms of stereotyped behaviour. *Journal of Applied Behaviour Analysis, 37,* 481-501. doi: 10.1901/jaba.2004.37-481

Shawler, L. A., & Miguel, C. F. (2015). The effects of motor and vocal response interruption and redirection on vocal stereotypy and appropriate vocalizations. *Behavioural Interventions,* 30, 112-134. https://doi.org/10.1002/bin.1407

Taylor, B. A., Hoch, H., & Weissman, M. (2005). The analysis and treatment of vocal stereotypy in a child with Autism. *Behavioural Interventions, 20,* 239-253. https://doi.org/10.1002/bin.200

Vollmer, T. R. (1994). The concept of automatic reinforcement: Implications for behavioral research in developmental disabilities. *Research in Developmental Disabilities, 15*(3), 187-207. https://doi.org/10.1016/0891-4222(94)90011-6

Wells, J., Collier, L., & Sheehey, P. H. (2016). Decreasing vocal stereotypy of a youth with Autism in a classroom setting. *Child and Family Behaviour Therapy, 38*(2), 164-174. https://doi.org/10.1080/07317107.2016.1172887

Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharczyk, S., Brock, M. E., Plavnick, J. B., Fleury, V. P., & .Schultz, T. R. (2014). *Evidence-based practices for children, youth, and young adults with autism spectrum disorder*. Chapel Hill: University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.

Wunderlich, K. L., & Vollmer, T. R. (2015). Data analysis of response interruption and redirection as a treatment for vocal stereotypy. *Journal of Applied Behaviour Analysis, 48*, 749-764. doi: 10.1002/jaba.227

Appendices

Appendix A: Figures

Figure 1

Python Script used to Calculate the Duration of Random Talking that is Yoked to the Previous Rird Session

```
import random




def gen_avg(expected_avg=5, n=13, a=3, b=9):
    while True:
        l = [random.randint(a, b) for i in range(n)]
        avg =reduce(lambda x, y: x + y, l) / len(l)


        if avg == expected_avg:
            return l


for i in range(100):
    print(gen_avg())
```

Figure 2

Percentage of Intervals with Vocal Stereotypy During Treatment Comparison for Eric (Pseudonym)

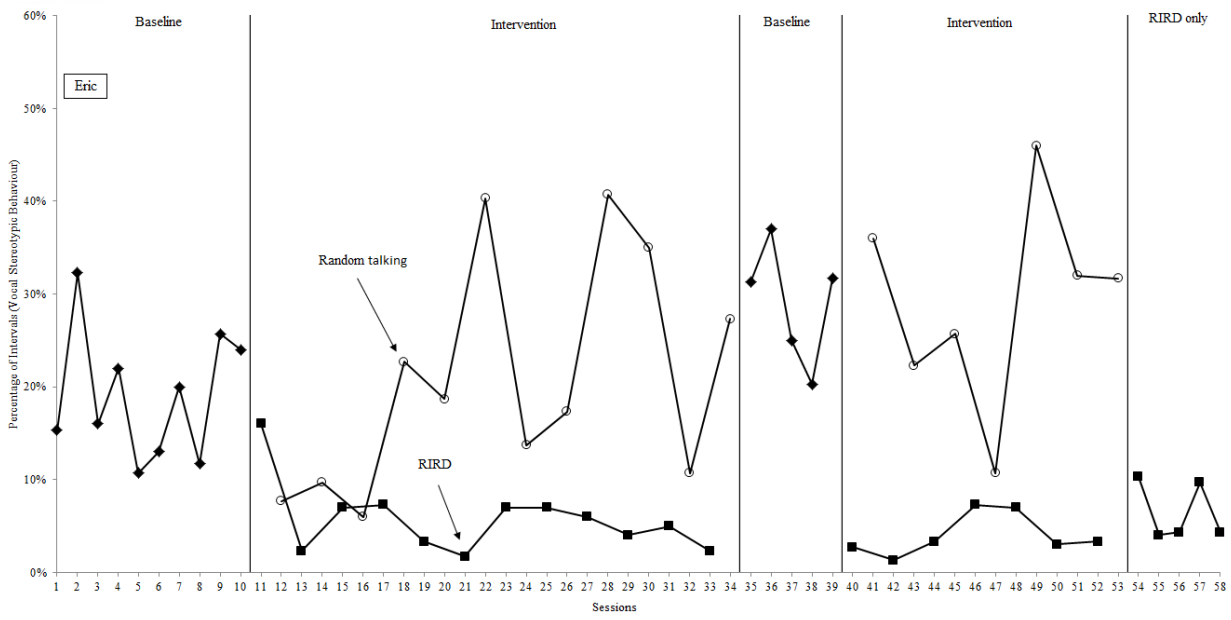
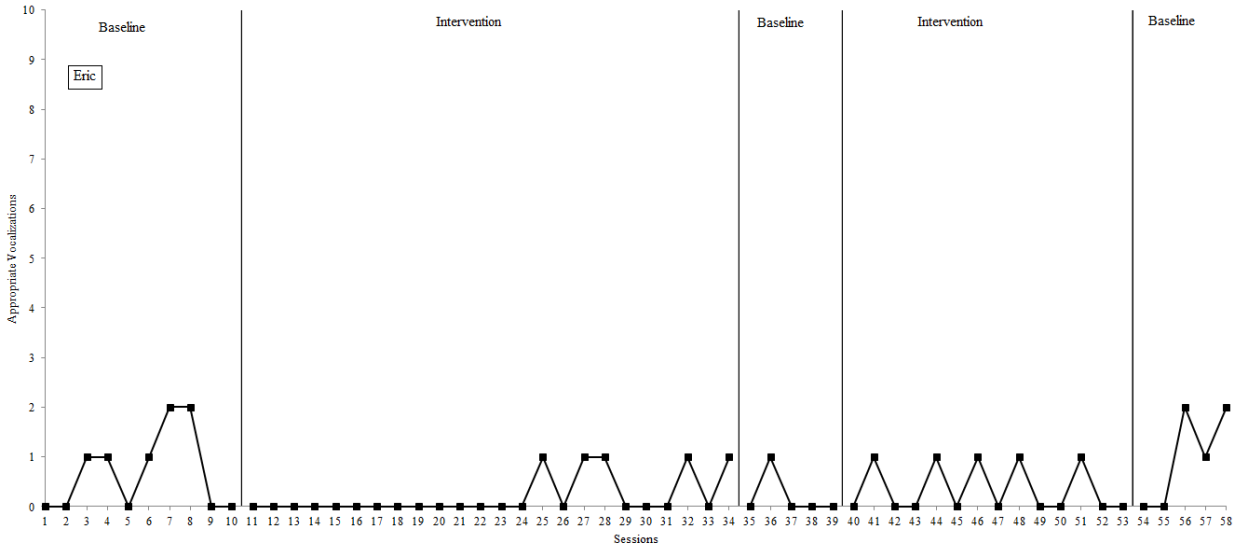


Figure 3

Frequency of Appropriate Vocalizations During Treatment Comparison for Eric (Pseudonym)



Appendix B

Phrases Used in Random-Talking Sessions

They said they would help us with the work
There is a lot of water in the ocean
I am going to need a hand with this
I would like to live in the country
I am going to read a book this weekend
It is not safe to play on the street
I was caught in the middle of the fight
You don't want to get lost in a desert
He lives on the bottom floor of his building
I like to drive at 100 km per hour
I had to measure the length of my couch
Why do we need to know what syllables are
Please don't sit on the edge of the hair
There are a lot of trees in the forest
A lot of people would love to be rich
It's not very safe at night time for kids
There was a lot of blood from her injury
We lost a lot of weight together this month
She went on with her work after the interruptions
Between you and me, she has had four divorces
His father would not permit his marrying of our daughter
The Earth travels round the Sun at great speed
You must read that letter, it contains important information
I am looking for a flat. Can you help me
Washington is hot in the summer, isn't it
They are twins, one is exactly like the other
How can I find him among so many people
Today I have a lot of work to do
He already has a car but wants another one.