10-5-2012

The effects of response interruption and redirection on language skills in children with vocal stereotypy

Tamara Leigh Perry

Follow this and additional works at: http://commons.emich.edu/theses

Part of the Clinical Psychology Commons

Recommended Citation

http://commons.emich.edu/theses/436

This Open Access Dissertation is brought to you for free and open access by the Master's Theses, and Doctoral Dissertations, and Graduate Capstone Projects at DigitalCommons@EMU. It has been accepted for inclusion in Master's Theses and Doctoral Dissertations by an authorized administrator of DigitalCommons@EMU. For more information, please contact lib-ir@emich.edu.
The Effects of Response Interruption and Redirection on Language Skills in Children with Vocal Stereotypy

by

Tamara L. Perry

Dissertation

Submitted to the Department of Psychology

Eastern Michigan University

in partial fulfillment for the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Clinical Psychology

Dissertation Committee:

James Todd, PhD, Chair

Renee Lajiness-O’Neill, PhD

Kenneth Rusiniak, PhD

Ivy Chong, PhD, BCBA-D

October 5, 2012

Ypsilanti, Michigan
Abstract

Vocal stereotypy is a common, skill-disruptive behavior in children with autism spectrum disorders (ASD). Response interruption and redirection (RIRD), the delivery of demands contingent on the occurrence of vocal stereotypy, is an intervention that is gaining empirical support for reducing vocal stereotypy and increasing appropriate language. However, little is known about the efficacy of RIRD when combined with early intensive behavioral intervention (EIBI), the recommended treatment approach for children with ASD, and its effects on the acquisition of language skills. Therefore, the purpose of this study was to evaluate the effects of RIRD when delivered during EIBI programming. Participants were six 4- and 5-year-old boys with ASD, assigned across three groups: experimental (EIBI + RIRD), clinical control (EIBI only), and traditional treatment control. Pre- and post-language skills were measured with the Preschool Language Scale-4 (PLS-4). Results indicated that significant group differences were noted in PLS-4 scores from pre- to post-testing, with large effects noted in the experimental group and small effects observed in the clinical control group. However, the participants in the traditional treatment control group outperformed all other participants across the PLS-4 language domains. EIBI + RIRD was successful in reducing vocal stereotypy for only one participant in the experimental group. Vocal stereotypy levels persisted in clinical control group participants during EIBI, confirming previous research that vocal stereotypy typically does not decrease without intervention. This study extends the literature on RIRD by demonstrating that it can be successfully integrated during EIBI programs; however, the necessary and sufficient conditions for RIRD to be effective have yet to be identified.
# Table of Contents

Abstract ................................................................................................................................. ii

Chapter 1: Introduction ........................................................................................................ 1  
  Distinguishing Between Vocal Stereotypy and Vocal Tics .................................................. 2  
  Development of Vocal Stereotypy ....................................................................................... 4  
  Functional Analysis ........................................................................................................... 7  
  Behavior Analytic Interventions for Vocal Stereotypy ....................................................... 10  
  Response Interruption and Redirection .............................................................................. 17  
  Using Language Acquisition Programs to Decrease Vocal Stereotypy ............................ 21

Chapter 2: Study Purpose .................................................................................................... 27  
  Hypotheses ....................................................................................................................... 27

Chapter 3: Method .............................................................................................................. 30  
  Participant Recruitment and Setting ................................................................................ 30  
  Participants ....................................................................................................................... 31  
  Assessment Measures ....................................................................................................... 34  
  Procedures for Experimental and Clinical Control Groups .............................................. 37  
  Data Collection and Interobserver Agreement ................................................................. 43  
  Data Analysis .................................................................................................................... 44

Chapter 4: Results .............................................................................................................. 46  
  Specific Aim I .................................................................................................................... 46  
  Specific Aim II ................................................................................................................... 50  
  Specific Aim III ................................................................................................................ 57

Chapter 5: Discussion ........................................................................................................ 59
Effectiveness of RIRD on Vocal Stereotypy during EIBI........................................59
Effects of RIRD on Standardized Scores of Language Skills ..................................63
The Relationships Between Language Skills, Developmental Level, and Vocal
Stereotypy ...................................................................................................................66
Future Research Directions ......................................................................................68
Limitations..................................................................................................................74
References..................................................................................................................77
Appendix A: Eastern Michigan University Human Subjects Approval Letter..............98
Appendix B: University of Nebraska Medical Center Human Subjects Approval Letter ......99
Appendix C: Informed Consent Form ........................................................................100
Appendix D: Steve’s Data Sheet.................................................................................103
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summary of Participant Characteristics</td>
</tr>
<tr>
<td>2</td>
<td>PLS-4 Results across Participants: Auditory Comprehension</td>
</tr>
<tr>
<td>3</td>
<td>PLS-4 Results across Participants: Expressive Communication</td>
</tr>
<tr>
<td>4</td>
<td>PLS-4 Results across Participants: Total Language Score</td>
</tr>
<tr>
<td>5</td>
<td>MSEL Results across Participants</td>
</tr>
<tr>
<td>6</td>
<td>Pearson Correlations between Vocal Stereoty, Language Skills, and Developmental Level of Functioning</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Outline of Study Procedures across Groups</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>Functional Analysis Results for Participants in the Experimental Group</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>Functional Analysis Results for Participants in the Clinical Control Group</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>Results of the Baseline during EIBI and EIBI plus RIRD Conditions for the</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Participants in the Experimental Group</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Results of the Extended Baseline Condition during EIBI for Participants in the</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Clinical Control Group</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Some children with autism spectrum disorders (ASD) and other developmental disabilities engage in vocal stereotypy, the persistent use of acontextual, non-purposeful, or repetitive vocalizations (Gardenier, MacDonald, & Green, 2004). Symptom presentation often varies in topography, duration, and severity but includes inappropriate noises (e.g., grunts, animal sounds), babbling, humming, and laughter in the absence of a humorous event. Vocal stereotypies can arise in the form of *echolalia*, the immediate or delayed imitation of previously heard sounds, words, or phrases emitted from other people or media sources (Stribling, Rae, & Dickerson, 2007), and *palilalia*, or delayed echolalia (Karmali, Greer, Nuzzolo-Gomez, Ross, & Rivera-Valdes, 2005). Vocal stereotypy is not unusual in children without ASD or other developmental disabilities during early language development. However, in these children, the behavior fades in frequency and persistence as functional language develops.

Although not a defining symptom of ASD, vocal stereotypy falls under the diagnostic criteria of “qualitative impairment in communication” in the DSM-IV-TR (American Psychiatric Association, 2000). Additional research on vocal stereotypy is needed because ASDs in general, and this symptom in particular, continue to be poorly understood. With the reported dramatic increase in the prevalence of ASD (Kim et al., 2011; Manning et al., 2011; Matson & Kozlowski, 2011) and the rise of controversial psychological and medical treatments without scientific basis (Jacobson, Foxx, & Mulick, 2005; Smith & Wick, 2008), the development of effective treatments for behavior problems is crucial. In contrast to other aberrant behaviors such as motor stereotypy (e.g., hand flapping, body rocking, finger twirling, etc.) or self-injury, in which physical response blocking can be used to temporarily
eliminate or moderate the problem, it is more difficult for clinicians to develop treatments to prevent vocal responding (Lanovaz & Sladeczek, 2012). Although engaging in vocal stereotypy does not place children in physical harm, existing data suggest that the presence of vocal stereotypy and other topographically-dissimilar forms of stereotypy is associated with negative social effects, in particular isolating children with ASD from others, including their typically developing peers (Smith & Van Houten, 1996), and disrupting the classroom environment (Athens, Vollmer, Sloman, & St. Peter-Pipkin, 2008). A child with severe vocal stereotypy might be so disruptive that he or she is rarely taken into many social settings. In light of the current push for inclusion in school settings, with the primary goal of developing social competence and success (Koegel, Koegel, Frea, & Freeden, 2001), it is becoming increasingly important to decrease problematic behaviors that differentiate children with ASD from their peers.

**Distinguishing between Vocal Stereotypy and Vocal Tics**

Individuals with ASD are at a higher risk than the general population for comorbid tic disorders (Baron-Cohen & Scahill, 1999; Canitano & Vivanti, 2007; Gadow, Devincent, Pomeroy, & Azizian, 2004; Gadow, Devincent, Pomeroy, & Azizian 2005). The prevalence of tic disorders in elementary school children is about 1% in non-clinical samples, and this figure increases in children who receive special education services, including those with ASD (Robertson, 2003). The presence of comorbid tic disorders can complicate the diagnostic situation as there is considerable overlap between vocal stereotypy and vocal tics, making it difficult for practitioners to discriminate between the two. Simple vocal tics include nonfunctional sounds, such as coughing, throat clearing, grunting, sniffing, snorting, chirping, barking, and other animal noises, whereas complex vocal tics include longer
utterances with words such as *coprolalia* (i.e., saying offensive or derogatory material), stuttering, immediate echolalia, and delayed echolalia (Woods et al., 2008). In the behavior analytic literature, the majority of these topographies are often classified as vocal stereotypy when they are irrelevant to the current context. However, echolalia is frequently cited as a common symptom in children with ASD and Tourette’s Disorder (Baron-Cohen & Scahill, 1999).

To meet DSM-IV-TR diagnostic criteria for a tic disorder, an individual must engage in many tics per day (the frequency is unspecified) for at least one year with an onset prior to 18 years of age (Woods et al., 2008). Furthermore, multiple motor and one or more vocal tics must be present in order to meet diagnostic criteria for Tourette’s Disorder (American Psychiatric Association, 2000). As noted, because they are topographically similar, differential diagnosis between vocal stereotypy and complex tics can be difficult. Aberrant vocal responses are considered more likely to be vocal stereotypy when the behavior appears rhythmic in nature, “more driven,” and self-soothing. Tics, in contrast, appear more “jerk-like” (Freeman, Soltanifar, & Baer, 2010), are thought to be involuntary, and occur in “temporal bouts” (American Psychiatric Association, 2000). Generally, they seem to appear suddenly. Coprolalia is not considered to be vocal stereotypy (Canitano & Vivanti, 2007). However, it is possible for true vocal stereotypy to coincidentally contain offensive words and ideas.

Besides similar topographies, vocal stereotypy and vocal tics have other shared features. Younger children are not typically aware of their vocal tics or vocal stereotypy. Therefore, it is harder for them than for adults to detect and actively suppress aberrant vocal responding. Tics occur more often when individuals are in a relaxed state, such as when they
are engaged in solitary activities in private areas (American Psychiatric Association, 2000). However, vocal stereotypy can also persist in the absence of social consequences for the majority of children with ASD (Cunningham & Schreibman, 2008), which suggests that this criterion may not be useful for distinguishing the two behaviors.

Gilbert (2006) provides several guidelines for discriminating between tics and stereotypy. Vocal tics do not occur when a person is speaking and instead are more likely to occur during pauses in speech or periods of silence. Tics are more likely to wax and wane across time and peak between 9 to 14 years of age, whereas vocal stereotypy levels remain more stable (Gilbert, 2006). Moreover, stereotypy emerges at a younger age than tics, typically before three years of age, with motor tics developing significantly earlier than vocal tics (Swain et al., 2007). Freeman and colleagues (2010) found a younger age of onset for stereotypic behavior in children with Stereotypic Movement Disorder (rather than ASD) compared to children with Tourette syndrome (mean age of onset = 17 months for stereotypy versus 5 years 10 months for tics). Finally, unlike tics, as noted earlier, vocal stereotypic behaviors are a normative part of human development and emerge early in life (e.g., babbling; Skinner, 1957, p. 41), suggesting that vocal stereotypy represents a retention of an infantile behavior pattern rather than the development of new and aberrant forms of responding. As such, it is important to understand why vocal stereotypy may develop and the benefits of engaging in vocal stereotypy as a prerequisite for language development.

**Development of Vocal Stereotypy**

In the behavior analytic literature, stereotypy is conceptualized as operant behavior, developed and maintained by the consequences that follow the response (Rapp & Vollmer, 2005). While the initial reinforcement for vocalizations is social, most authors presume
stereotypy subsequently is maintained primarily by *automatic reinforcement*—reinforcement that is not socially mediated and arises from the property of the behavior itself (Vaughn & Michael, 1982). In terms of vocal stereotypy, automatic reinforcement refers to the auditory stimulation provided by the act of producing sounds. Vollmer (1994) suggested two reasons why automatically-reinforced behaviors are challenging to reduce. First, some aspects of the reinforcers are inaccessible and cannot be objectively scored and directly manipulated. For example, vocal stereotypy may be maintained by sensations caused by the vibration of the vocal cords, which cannot be directly manipulated within an intervention (Lanovaz & Sladeczek, 2012). Likewise, what is heard by the person himself or herself will likely differ considerably from what is heard by outside observers, especially if the stereotyped vocalizations are of low intensity. Furthermore, well-established treatments for automatically-maintained problem behavior have not yet been developed. Instead, effective treatments for eliminating behavior disorders have focused on socially mediated functions, such as attention and escape from demands.

Lovaas, Newsom, and Hickman (1987) developed the “Perceptual Reinforcement Hypothesis” to describe the role of automatic reinforcement in the maintenance of various topographies of stereotypy, including (but not limited to) vocal stereotypy, hand flapping, body rocking, and lining up objects. This hypothesis states that “self-stimulatory” behaviors are maintained by the immediate, potent sensations (i.e., “perceptual reinforcers”) produced by engaging in stereotypy. This interpretation provides a rationale for why other forms of stereotypy sometimes emerge or increase if one response is blocked or disrupted; if powerful perceptual reinforcers control stereotypy, then children will seek out other alternatives to contact sensory reinforcement.
Some repetitive vocalizations that are maintained by automatic reinforcement can be adaptive, however. Babbling and vocal play are important parts of the language acquisition process, as vocal play is a necessary precursor to developing vocal imitation skills (Sundberg & Partington, 1998). When children with ASD are entirely nonvocal, babbling and approximations to words must be taught. Lovaas, Koegel, Simmons, and Long (1973) found that children with ASD who engaged in echolalia had better prognoses than children who made no sounds because they already had vocal imitation skills within their behavioral repertoire. Although parental praise may be a first step in strengthening babbling, automatic reinforcement explains why infants babble excessively even when parents are not immediately present or within close proximity (Bijou & Baer, 1965; Sundberg & Michael, 2001). That is, the sounds produced by babbling may function as conditioned reinforcers.

A limitation of the perceptual reinforcement hypothesis is that a small number of treatment studies have identified social and communicative functions for vocal stereotypy (e.g., Frea & Hughes, 1997; Rehfeldt & Chambers, 2003). Although attention- or escape-maintained vocal stereotypy has been cited less frequently than cases of automatically-reinforced stereotypy, function is idiosyncratic and varies across children (Turner, 1999). Echolalia, in particular, has been shown to increase when a child does not know the answer to a question (Carr, Schreibman, & Lovaas, 1975; Schreibman & Carr, 1978). For instance, Stribling and colleagues (2007) identified possible communicative functions of echolalia in a 16-year-old girl with ASD. Through an analysis of videotaped segments across various classroom activities, the authors observed that immediate echolalia occurred only after someone had spoken to the participant directly and did not occur when others were engaging in conversation. Stribling et al. (2007) hypothesized that the participant used echolalia as a
rudimentary means to acknowledge that (1) she understood that the initiation was directed towards her and (2) she was expected to provide an answer. In addition, some children with ASD may engage in echolalia to make requests or participate in conversational exchanges (McEvoy, Loveland, & Landry, 1988; Prizant & Duchan, 1981). It is also certainly possible that stereotypies might persist because of intermittent reinforcement. A response reinforced once every hundredth occurrence, on average, would likely appear on casual observation, to not have an external reinforcer at all. Given that vocal stereotypy may serve multiple idiosyncratic functions for a child with ASD, conducting a functional analysis is a customary procedure in both clinical and research settings. We should not be surprised, however, if determining the function is difficult.

**Functional Analysis**

Functional analysis is considered the “gold standard” for determining the consequences of behavior in individuals with ASD and other developmental disabilities (Herzinger & Campbell, 2007). Clinicians can use experimental methods to test hypotheses about possible functions to rule in or rule out plausible explanations for the maintenance of responding. Iwata, Dorsey, Slifer, Bauman, and Richman (1994) first described the general model of the functional analysis procedure, which was developed as an extension of well-understood principles of the experimental analysis of behavior (Sidman, 1960) to the problem of identifying controlling variables of destructive behavior problems in applied settings. That is, a functional analysis is an experiment designed to assess and test for the effects of various likely reinforcement contingencies for the target behavior.

Functional analysis procedures have been adapted to assess the function of vocal stereotypy and other forms of aberrant vocal behaviors—including bizarre speech displayed
by patients with schizophrenia and bipolar disorder (Lancaster et al., 2004) and vocal tics (Watson & Sterling, 1998). The standard functional analysis protocol in these instances typically includes four conditions: attention, demand, alone, and play (Iwata & Dozier, 2008). The attention condition is used to determine whether the child engages in vocal stereotypy in order to receive the attention from others, typically an adult. During this test condition, the adult sits next to the child and pretends to read a book or magazine. If the child engages in vocal stereotypy, the adult turns to the child and delivers brief reprimands, such as, “Stop making so much noise. You need to be quiet and play with your toys while I am reading,” and then resumes pretending to read the materials until another instance of vocal stereotypy occurs. The demand condition tests the negative reinforcement contingency by investigating if the child engages in vocal stereotypy during academic or similar work in order to “escape” (break) from tasks. In the alone condition, the child is in a room by himself or herself to simulate a sensory-deprived environment to determine if vocal stereotypy is maintained by automatic reinforcement. An ignore condition is frequently substituted for the alone condition, in which a therapist is present in the room with the child but does not interact with the child in any manner. Last, the play condition serves as the control, at which time the child has free access to preferred toys, child-directed activities with no delivery of demands, and the therapist’s noncontingent attention.

Upon completion of the functional analysis, data are analyzed to identify the function of vocal stereotypy based on patterns of responding. Generally, if rates of responding are significantly elevated in one condition than in other conditions, the contingency operating in that condition is inferred to be the maintaining event. For instance, if the child engages in high rates of vocal stereotypy during the attention condition only, clinicians would assume
that vocal stereotypy is maintained by social attention. In another example, a child may engage in higher levels of vocal stereotypy during the demand condition than other conditions, suggesting that he or she engages in vocal stereotypy to escape from nonpreferred or less preferred tasks. Because the play condition serves as the control condition, it is expected that the child will not engage in high rates of vocal stereotypy during this arrangement. It is possible to have stereotypy maintained across all conditions or to not appear with sufficient frequency to identify a function.

Two response patterns are indicative of an automatic function (LeBlanc, Patel, & Carr, 2000). In the first case, differentially higher rates of vocal stereotypy are observed in the alone or ignore condition, and zero to low rates are observed across the remaining conditions. Alternatively, the child engages in high rates of vocal stereotypy across all conditions, a pattern often referred to as “undifferentiated.” Undifferentiated patterns of responding could indicate that vocal stereotypy is maintained by multiple functions, including those that are socially and nonsocially mediated. For example, Kennedy, Meyer, Knowles, and Shukla (2000) found that multiple reinforcers were responsible for behavioral maintenance for various forms of stereotypy in five children with ASD. Functional analyses revealed that both attention and escape from demands maintained stereotypy in two participants, and automatic reinforcement and escape were identified as the functions for another child. The remaining two participants engaged in high levels of stereotypy across all conditions. Although based on these data patterns, it appeared that stereotypy may have been reinforced by automatic internal stimulation for these two children; teaching them to make requests for each reinforcer was successful in reducing stereotypy, thereby supporting the hypothesis that stereotypy served multiple, simultaneous functions.
When results from the functional analysis are judged to be inconclusive, some authors advise conducting multiple, consecutive alone or ignore conditions to confirm that the child engages in consistently high rates of vocal stereotypy in the absence of social contingencies (e.g., Iwata & Dozier, 2008). Vollmer, Marcus, Ringdahl, and Roane (1995) recommended altering the experimental design when data patterns are ambiguous. For instance, clinicians could implement a reversal design, in which several sessions of the same condition are conducted in successive order until a stable pattern emerges before reinitiating another functional analysis condition. Pairwise designs, the combination of multielement and reversal designs, are also used on a regular basis when the function of behavior is unclear (Iwata & Dozier, 2008). Data gathered from the functional analysis are then used to guide treatment decisions.

**Behavior Analytic Interventions for Vocal Stereotypy**

Vocal stereotypy exhibited by individuals with ASD has been reduced or eliminated by a variety of techniques, including noncontingent reinforcement (e.g., Ahearn, Clark, DeBar, & Florentino, 2005), differential reinforcement of other behavior (e.g., Mancina, Tankersley, Kamps, Kravits, & Parrett, 2000), differential reinforcement of alternative behavior (e.g., Frea & Hughes, 1997; Rehfeldt & Chambers, 2003), punishment of stereotyped vocalizations (e.g., Rapp et al., 2009), and discrete-trial training (e.g., Dib & Sturmey, 2007; Schreibman & Carr, 1978). Behavior analytic interventions are typically categorized into antecedent and consequent manipulations. Antecedent interventions include manipulating the environment to occasion and increase desirable behavior by increasing the availability of reinforcement or altering the effectiveness of reinforcers, whereas consequent
interventions involve manipulations of the consequences that maintain a particular problem behavior (Cooper, Heron, & Heward, 2007).

One type of antecedent intervention that has been well researched in the stereotypy literature is the use of matched stimuli, items that are similar to the overt sensory input produced by engaging in stereotypy, such as auditory stimulation (Rapp, 2007). Matched items often are delivered via noncontingent reinforcement (NCR) schedules, the response-independent delivery of a reinforcer on a time-based schedule (Carr et al., 2000). There are several demonstrations of the effectiveness of the noncontingent delivery of matched stimuli to reduce motor stereotypy (e.g., Higbee, Chang, & Endicott, 2005; Piazza, Adelinis, Hanley, Goh, & Delia, 2000; Rapp, 2006; Sidener, Carr, & Firth, 2005). However, not all efforts have been successful. Rapp (2006) found that three boys with ASD engaged in varying levels of vocal and motor stereotypy when presented with auditory and visual stimulation from a television. For one participant, increased vocalizations were observed when the sound was activated, while vocal stereotypy decreased temporarily when the television was muted; this reduction in behavior is not explained by the premise of matched stimuli because the television provided visual rather than auditory stimulation during this arrangement. In a further investigation of matched stimuli as treatment for vocal stereotypy, Rapp (2007) found that while the noncontingent delivery of matched toys reduced vocal stereotypy in one boy with ASD, some matched toys actually increased repetitive vocalizations in the other participant. These findings suggest that the suppressive effects of matched stimulation may be idiosyncratic across participants and treatments (because maintaining conditions likely differ) and treatments should be individually tailored based on information from a functional analysis.
Alternatively, Ahearn, Clark, DeBar, and Florentino (2005) supported the use of unmatched, preferred reinforcers within NCR schedules across two children with autism and comorbid intellectual disability. Results indicated that access to a matched toy decreased rates of vocal stereotypy by approximately 50% in both participants but significantly lower rates were observed during access to the preferred, unmatched toy. These data suggest that preference level for a reinforcer may be responsible for response suppression than similarity to sensory stimulation provided by vocal stereotypy. That is, while stereotypy might provide a reinforcing experience, other reinforcers are likely to be more powerful. Furthermore, duration of engagement, as a measure of preference, may not always be a useful predictor of treatment effectiveness.

Consequent interventions are more commonly used in the applied literature for treatment of vocal stereotypy and encompass many procedures, including differential reinforcement, extinction, and punishment. Differential reinforcement of other behavior (DRO), the reinforcement of zero rates of the target behavior, is one type of consequent intervention that is frequently used within applied settings (Repp, Deitz, & Speir, 1974; Reynolds, 1961). During this procedure, the practitioner delivers a reinforcer only after a specified interval in which the target behavior does not occur. Differential reinforcement procedures are often combined with extinction, an intervention in which the reinforcement contingency that maintains responding is discontinued via the withholding of that reinforcer (Sidman, 1960, p. 402). In the case of vocal stereotypy maintained by attention, for example, the therapist would apply extinction by withholding comments or eye contact to the child contingent on instances of vocal stereotypy.
DRO has also been applied to vocal stereotypy maintained by sensory consequences. Taylor, Hoch, and Weissman (2005) applied DRO to automatically-reinforced vocal stereotypy in a 4-year-old girl with ASD. The first phase consisted of an antecedent analysis with matched toys that produced auditory stimulation via music and sounds, with the goal of identifying stimuli that would compete with vocal stereotypy. To ensure that response suppression was due to the competition of auditory input rather than other qualities of the toys, the rate of stereotypy was also measured when the batteries were removed from the toys. Results from this antecedent analysis confirmed that vocal stereotypy decreased to near zero levels during access to matched toys but persisted when the toys were inoperable.

Matched toys were then included in a DRO contingency, in which the participant sat with the experimenter at a table with less preferred toys and was delivered the instruction, “If you play quietly, when the timer rings, you can play with the music toys.” A 1-minute timer was activated, and if the child did not engage in vocal stereotypy, she earned 30-second access to the auditory toys, while contingent on the engagement of vocal stereotypy, the timer was reset. DRO intervals were gradually increased to 10 minutes and treatment gains successfully generalized to her classroom setting.

Similar to DRO, differential reinforcement of alternative behavior (DRA) is a procedure in which appropriate replacement behaviors are reinforced in the absence of the target behavior (Vollmer & Iwata, 1992). Current DRA literature for vocal stereotypy focuses only on behaviors that are maintained by social consequences; that is, to date, no studies with DRA have been applied to automatically-maintained topographies of vocal stereotypy. Frea and Hughes (1997) decreased perseverative speech and inappropriate laughter in two adolescents with intellectual disabilities using a combination of DRA and
extinction procedures. Results of a functional analysis determined that vocal stereotypy was maintained by social contingencies in the forms of attention and escape for both participants. Functionally-equivalent replacement responses for stereotypy were trained, such as “excuse me” for attention-maintained perseverative speech and “I don’t know” for escape-maintained inappropriate laughter. Contingent on emitting appropriate communication, participants received either attention or escape from work tasks, while all instances of vocal stereotypy were ignored.

Rehfeldt and Chambers (2003) reduced perseverative speech maintained by attention in a 23-year-old man with ASD using a reversal design. Perseverative speech was defined as repetitive statements regarding circumscribed interests, including sirens, medical appointments, and coughing. In contrast, appropriate speech was considered any initiation or response that was unrelated to the target topics. DRA plus extinction consisted of delivering 5 seconds of attention and eye contact contingent on appropriate verbalizations and ignoring perseverative speech by removing eye contact and decreasing proximity to the participant. During the baseline sessions, the participant engaged in approximately 30 perseverative statements within a 10-minute period with few appropriate verbalizations. Across treatment sessions, he emitted approximately 30 to 40 appropriate vocal responses and 10 or fewer perseverative statements per session. A study limitation was that reductions in perseverative statements were not generalized to the participant’s natural environment.

Punishment procedures decrease the probability of the future occurrence of the target behavior through the contingent presentation of an aversive stimulus or removal of a reinforcing one (Skinner, 1953, p. 185). Technically, the nature of the events does not need to be defined in advance and is inferred from their effects. Skinner (1953, p. 240) noted that
punishment is sometimes successful in attenuating the effects of automatic reinforcement. At that time, however, punishment was considered an indirect procedure in which the direct effects of aversive stimulation produced behavior that interfered with the ongoing operant response, as in conditioned suppression (Estes & Skinner, 1941). But application of punishment is controversial. Although punishment procedures can lead to rapid deceleration in response rates (e.g., Azrin, 1960; Vorndran & Lerman, 2006), some consider the use of punishment techniques a last resort, especially when positive punishment—the onset of aversive stimulation contingent on behavior—is used (e.g., Beare, Severson, & Brandt, 2004; DiGennaro Reed & Lovett, 2008). Moreover, unlike reinforcement-based procedures such as DRA, most punishment strategies do not explicitly increase desired behaviors or teach appropriate replacement responses (Herzinger & Campbell, 2007). Despite these ethical concerns, Lerman and Vorndran (2002) emphasize the necessity of continuing to conduct research on punishment procedures in order to understand how to reduce automatically-maintained behaviors more efficiently and effectively. In fact, the use of punishment may be the most ethical treatment choice if alternative, less intrusive methods are unsuccessful or unavailable (Lerman & Vorndran, 2002). To prevent some of the unwanted effects associated with punishment (e.g., increased aggressive behavior), some researchers have combined punishment procedures with other established response-suppression techniques.

Research has focused on using less aversive and restrictive forms of punishment in combination with NCR (e.g., Athens, Vollmer, Sloman, & St. Peter-Pipkin, 2008). Response cost, the contingent loss of reinforcers, sometimes called “negative punishment,” is one such alternative because the child is not harmed and has the opportunity to earn access to the reinforcer once the inappropriate behavior ceases (Cooper et al, 2007). Falcomata, Roane,
Hovanetz, Kettering, and Keeney (2004) compared NCR with a matched stimulus to NCR plus response cost within a reversal design as treatment for noncontextual vocalizations in an 18-year-old man. A preference assessment was conducted to select the matched reinforcer, a radio, which was incorporated into the NCR schedule. In the NCR condition, the participant had continuous access to the radio. During the combined NCR and response-cost condition, the participant had continuous access to the radio but contingent on vocal stereotypy, the radio was silenced for 5 seconds. Results revealed that vocal stereotypy decreased by only 40% during NCR conditions but reduced to levels near zero during the NCR plus response cost arrangement. Similar to the procedures used by Falcomata et al. (2004), Athens and colleagues (2008) developed a treatment package of NCR, contingent demands, and response cost to reduce vocal stereotypy in an 11-year-old boy diagnosed with ASD and comorbid Down syndrome. Contingent on vocal stereotypy, the boy lost access to preferred toys and was delivered contingent demands, in which he was required to provide a vocal response to academic tasks. Findings indicated that the treatment package was successful in decreasing vocal stereotypy even when noncontingent attention was removed, and the therapist was faded from the room.

Recent studies have used less restrictive forms of punishment procedures that teach language skills while suppressing vocal stereotypy. For example, Karmali and colleagues (2005) found tact correction to be effective in reducing delayed echolalia in five preschoolers with ASD. A tact is a verbal operant that includes the behaviors of labeling, naming, and describing environmental stimuli and is evoked by a nonverbal discriminative stimulus (Skinner, 1957, pp. 81-83). Tact behavior increases in children through the attention of a verbal audience, in which it is maintained by nonspecific reinforcement, such as a
generalized social reinforcer (e.g., verbal praise, adult attention). The authors developed tact correction as a form of treatment with the supposition that if participants learned to provide tacts about their environment, tact behavior may become automatically reinforcing via adult praise and eventually replace vocal stereotypy. Contingent on vocal stereotypy, therapists engaged in tacts relevant to the current context until the child stopped emitting vocal stereotypy.

**Response Interruption and Redirection**

Another technique that incorporates vocal language tasks is response interruption and redirection (RIRD), which will be the focus of the current study. RIRD, a systematic procedure to apply response blocking to vocal behavior, was originally developed by Ahearn, Clark, MacDonald, and Chung (2007) to decrease automatically-maintained vocal stereotypy. With regard to consequence-based interventions for vocal stereotypy in children with ASD, RIRD is a behavior-reduction procedure with the most empirical evidence (Lanovaz & Sladeczek, 2012). This method was based on response blocking techniques used for other types of aberrant behaviors for individuals with intellectual disabilities (e.g., Rincover, 1978). Redirection procedures were also added to eliminate the child from continuing to engage in vocal stereotypy. Both components have been considered necessary for treatment success because response blocking alone has been associated with negative side effects, including aggressive behavior and emotional outbursts (Hagopian & Adelinis, 2001). In the Ahearn et al. study, contingent on the occurrence of vocal stereotypy, therapists delivered vocal tasks until the child repeated three consecutive verbal utterances or answered social questions (e.g., “What’s your name?”) correctly without displaying further vocal stereotypy. Ahearn et al. reported clinically significant reductions in vocal stereotypy in four children.
with ASD compared to baseline levels. Additionally, follow-up data indicated that vocal stereotypy occurred at lower rates in the natural environment during post-intervention probes.

Recent studies have replicated the findings of the Ahearn et al. (2007) study. For example, Liu-Gitz and Banda (2010) demonstrated effective abatement of vocal stereotypic behavior maintained by automatic reinforcement in a student with ASD within a reversal design. The participant’s teacher conducted all RIRD sessions within his classroom, demonstrating that RIRD can be efficient and practical to implement with children who are not in 1:1 academic settings. Miguel, Clark, Tereshko, and Ahearn (2009) found that RIRD was more effective than pharmacotherapy (i.e., sertraline) in reducing vocal stereotypy in a preschooer with ASD. The single and combined effects of RIRD and sertraline were evaluated within a ABABC reversal design, in which vocal stereotypy levels were measured during sertraline only (A), RIRD + sertraline (B), and RIRD only (C) conditions. Results indicated that sertraline had no added effects because no increases in vocal stereotypy were observed after the medication was faded systematically and vocal stereotypy levels were similar across RIRD + sertraline and RIRD only conditions, demonstrating that RIRD was the active agent of change in the sertraline plus RIRD package. Of note, the participant engaged in higher rates of appropriate manding when RIRD was in place.

Casella, Sidener, Sidener, and Progar (2011) replicated the results of Ahearn et al. (2007) with two boys with ASD and vocal stereotypy and extended their findings by demonstrating treatment effectiveness with motor RIRD tasks. Contingent on vocal stereotypy, the experimenter delivered motor demands in the form of one-step commands (e.g., “touch head”). Results showed that RIRD was successful in reducing vocal stereotypy, despite that RIRD tasks did not match the topography of stereotypy. That is, the delivery of
motor tasks was sufficient in interrupting and redirecting vocal stereotypy without the use of an incompatible response. One potential concern with RIRD is that the frequency of vocal stereotypy may increase undesirably when the treatment is not being implemented because the individual had been previously deprived of the stimulation provided by vocal stereotypy during RIRD. To examine this hypothesis, Schumacher and Rapp (2011) evaluated the effects of RIRD within a three-component multiple-schedule for two participants. The baseline condition consisted of a no-interaction sequence, in which the therapist was present in the room with the child but did not interact with the child and instead pretended to read written materials, similar to an ignore condition of a functional analysis. No materials were on the walls during the no-interaction sessions as a signal for participants that no consequences would be applied to vocal stereotypy. During RIRD sequences, the first and third components were conducted identically to the no-interaction condition. RIRD tasks were delivered during the second component, which was signaled by the therapist applying a red piece of construction paper on the wall. RIRD tasks consisted of previously acquired vocal demands in the forms of tacts (e.g., ongoing actions, body parts) and intraverbals (e.g., animal sounds, personal information, and categories of objects). Findings indicated that vocal stereotypy levels were similar across the first and third components of the RIRD sequence, in which RIRD was not in effect, demonstrating that vocal stereotypy did not increase upon the removal of RIRD. In addition, RIRD resulted in immediate decreases in vocal stereotypy levels across both participants.

Most likely, RIRD is an effective procedure for several reasons. Conceptually, it is similar to positive practice, a type of overcorrection technique in which the individual is
required to engage in the correct response repeatedly for a certain number of times (Foxx & Azrin, 1973). Furthermore, an incompatible response is often used, decreasing the likelihood of vocal stereotypy because it is impossible for the child to make correct vocal utterances and emit vocal stereotypy simultaneously. Although RIRD may be viewed as a sensory extinction procedure, in that the automatic reinforcement contingency that maintains responding is temporarily blocked, the process of punishment is responsible for behavioral change because the frequent presentation of demands is an aversive contingent event (Ahrens, Lerman, Kodak, Worsdell, & Keegan, 2011).

To evaluate the operative mechanism for RIRD, Ahrens et al. (2011) compared the effects of vocal and motor RIRD on both vocal and motor stereotypic responses. The rationale for this comparison was that if RIRD tasks served as punishers, vocal stereotypy would decrease regardless of the form of task, whereas if the operative mechanism for RIRD was sensory extinction, vocal RIRD would have greater suppressive effects on vocal stereotypy than motor RIRD. Results indicated that both forms of RIRD were successful in decreasing vocal stereotypy, irrespective of the topography of stereotypy, thereby supporting the interpretation that RIRD is a punishment procedure. The authors further investigated the punishment hypothesis with a 4-year-old boy with ASD by manipulating the procedural integrity of RIRD. Specifically, vocal stereotypy levels were compared across a baseline with no programmatic consequences for vocal stereotypy and three fading conditions (i.e., interrupting 50%, 25%, and 10% of vocal stereotypic responses) within a reversal design. This arrangement was chosen to observe if (1) vocal stereotypy decreased as treatment integrity increased, which would demonstrate that punishment was the operative mechanism, or (2) vocal stereotypy increased as treatment integrity increased, which would indicate
sensory extinction because the automatic reinforcement schedule would be thinned. Results revealed that vocal stereotypy decreased to low levels during the 50% treatment-integrity condition, and levels were slightly higher than baseline levels in the 25% treatment-integrity condition. In addition, vocal stereotypy increased to baseline levels when RIRD was implemented for every 10th response (i.e., 10% treatment-integrity condition). These findings confirmed the punishment interpretation because vocal stereotypic responses abated as more responses were followed by a consequence.

RIRD has an inherent teaching component, in which previously acquired vocal tasks are practiced (e.g., answering the question, “How old are you?”). The repetition of these tasks increases the likelihood that the child will maintain these responses over time and may increase the spontaneous use of vocal language. In fact, although appropriate vocalizations were not specifically targeted during treatment, Ahearn et al. (2007) identified concomitant gains in manding behavior in three participants during RIRD conditions. Despite these gains, a limitation to this study was that the participants were not required to engage in other activities during the implementation of treatment (e.g., academic work). To increase the generality of this procedure, further research should focus on embedding RIRD within existing skill acquisition programs for children with ASD.

**Using Language Acquisition Programs to Decrease Vocal Stereotypy**

As previously stated, a primary purpose for developing effective treatments is to reduce inappropriate behaviors such as vocal stereotypy that interfere with the acquisition of language and social skills in young children with ASD. Early intensive behavioral intervention (EIBI) is a well-established intervention for young children with ASD that targets the core symptoms of autism (Eikeseth, 2009; Reichow, 2012; Rogers & Vismara,
EIBI uses repetition and sequenced one-to-one instruction to teach appropriate social, communicative, and academic skills from an applied behavior analysis (ABA) framework (Smith, Mozingo, Mruzek, & Zarcone, 2007). ABA, the application of Skinner’s principles of behavior analysis to human behavior, is considerably effective when the principal goal of intensive intervention is to promote socio-communicative skills (Sundberg & Michael, 2001). Early intervention is recommended because the prognosis for children who have not developed functional communication skills after age five is less favorable (Pickett, Pullara, O’Grady, & Gordon, 2009).

Numerous studies have reported clinically significant increases in standard scores on tests measuring cognitive and developmental functioning following EIBI (e.g., Anan, Warner, McGillivary, Chong, & Hines, 2008; Andersen, Avery, DiPietro, Edwards, & Christian, 1987; Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovaas, 1987; Lovaas et al., 1973; McEachin, Smith, & Lovaas, 1993; Sallows & Graupner, 2005; Smith, 1999; Smith, Groen, & Wynn, 2000; Smith et al., 2010). In the seminal UCLA study, Lovaas (1987) reported average gains of 20 IQ points in 19 children following two years of intensive behavioral intervention with one-to-one structured instruction, compared with little to no IQ gains in children who received only 10 or fewer hours of weekly intervention. Additionally, some studies have documented increases in participant standard scores on adaptive measures following EIBI as well (e.g., Magiati, Charman, & Howlin, 2007; Sallows & Graupner, 2005; Smith et al., 2010; Weiss, 1999).

Recent research also validates the use of EIBI over traditional treatment modalities (e.g., Cohen, Amerine-Dickens, & Smith, 2006; Eikeseth, Smith, Jahr, & Eldevik, 2007; Remington et al., 2007). For instance, Howard, Sparkman, Cohen, Green, and Stanislaw
(2005) demonstrated the efficacy of EIBI over eclectic treatments (i.e., a combination of the Picture Exchange Communication System, sensory integration, and social activities used in mainstream classrooms) and center-based special education services. Findings showed that children who received intensive behavioral treatment (n = 29) scored significantly higher on standardized tests measuring IQ, language abilities, and adaptive skills than children who participated in the alternative interventions after 14 months of treatment (n = 16).

Sundberg and Partington (1998) developed a widely-used treatment manual for young children with ASD based on Skinner’s (1957) analysis of verbal behavior. To apply the curriculum within EIBI, behavior analysts tailor programs based on the child’s current repertoire to teach a variety of skills, such as visual performance (e.g., puzzles, block design), receptive language, expressive language, adaptive, and fine-motor skills. Each program consists of breaking down target skills into smaller components and then teaching each component in a discrete step with prompting strategies and reinforcement (Gould, Dixon, Najdowski, Smith, & Tarbox, 2011). Verbal operant training is a particular focus, including teaching mands (e.g., making requests when items are not present, using mand frames such as, “I would like” and “May I have ____?”), tacts (e.g., labeling and describing environmental stimuli and people), and echoics (i.e., vocal imitation of sounds, words, etc.). In addition, intraverbals skills are targeted to teach children how to answer many questions. Intraverbals are evoked by a verbal stimulus, do not have one-to-one correspondence, and are maintained by generalized conditioned reinforcers (Cooper et al., 2007; Skinner, 1957). Examples of common intraverbal skills taught in EIBI programs include answering questions related to personal information (e.g., name of child’s school, age, birthday, telephone
VOCAL STEREOTYPY

number, etc.), features, functions, and classes of items, object associations (e.g., “Shoes and socks”), and filling in the blanks to children’s songs (e.g., “Itsy, bitsy, spider”).

Although the primary goals for EIBI are related to increasing functional skills related to communication and socialization, clinicians often apply treatment strategies for reducing vocal stereotypy within EIBI. Reducing problem behavior and developing instructional control are often prerequisites to effective learning. Kodak, Fisher, Clements, Paden, and Dickes (2011) identified several variables that influenced correct responding during receptive language tasks (i.e., pointing to pictures in an array of three or four items following the discriminative stimulus, “Point to____”), including problems with inattention, motivation, and responding to controlling prompts. These instructional variables were idiosyncratic across children and interventions targeted to decrease barriers to learning identified from functional assessment results were more effective than alternative, randomly-selected instructional procedures (e.g., prompting strategies). While beyond the scope of the Kodak et al. study, vocal stereotypy could be considered a barrier to learning for children who engage in high levels of this behavior. Data showing that suppression of vocal stereotypy are correlated with increased skill acquisition supports this rationale. For example, Lovaas et al. (1973) observed systematic increases in appropriate vocalizations once echolalia was eliminated across participants with ASD. Therefore, several earlier studies have used EIBI to replace vocal stereotypy with appropriate language.

Carr, Schreibman, and Lovaas (1975) observed that children with ASD were more likely to engage in vocal stereotypy in response to novel questions than to familiar, practiced questions. Consistent with this observation, Schreibman and Carr (1978) taught two children to provide appropriate answers to social questions to eliminate echolalia in a multiple-
VOCAL STEREOTYPY

baseline across participants design. Training sessions were used to teach the children to say, “I don’t know” in response to previously repeated “what,” “how” and “who” questions via a verbal prompting procedure. During the second phase of the study, training questions were interspersed with questions that the children had a history of answering correctly without immediate echolalia (e.g., “What is your name?”) to test if the children were able to discriminate between unknown and mastered questions. Generalization probes were then conducted with novel questions. Results indicated that both participants were able to respond appropriately without engaging in vocal stereotypy to training and novel questions for at least 10 consecutive sessions. Treatment gains were maintained one month later for both participants.

McMorrow and Foxx (1986) replaced echolalia with appropriate question answering in a 21-year-old man with ASD and comorbid intellectual disability. Treatment procedures included teaching the participant to remain silent when a question was presented and subsequently to respond correctly based on naturally-occurring discriminative stimuli. Because the participant was able to tact pictures prior to intervention, response cards were used as visual prompts to facilitate correct responding. Differential reinforcement was delivered for correct responses without echolalia. More recently, Foxx, Schreck, Garito, Smith, and Weisenberger (2004) extended the findings of McMorrow and Foxx (1986) by teaching two children to respond appropriately to factual questions about familiar objects and animals that they had been previously taught to tact. Children were trained to pause while each question was asked and then point to the picture card of the correct answer. Picture cards were gradually faded and the children were required to answer the question with a vocal response.
Although sometimes conceptualized as a positive punishment procedure (Ahrens et al., 2011), RIRD is consistent with EIBI strategies and provides children with opportunities to practice vocal language skills. RIRD also has been demonstrated to be effective in reducing vocal stereotypy when used as an adjunctive procedure to verbal operant training. Specifically, Colon, Ahearn, Clark, and Malasky (2012) evaluated the independent effects of teaching children tacts and the combined effects of tact training and RIRD based on the premise that teaching vocal communication skills that contact social reinforcement may decrease vocal stereotypy via response competition. Findings indicated that verbal operant training alone was insufficient in reducing vocal stereotypy to clinically acceptable levels in two out of three participants. Therefore, RIRD was introduced as a single intervention. Results indicated that RIRD was only slightly more effective in suppressing vocal stereotypy for one participant relative to tact training; for this participant, RIRD was then implemented in combination with tact training, and further suppression was observed compared to RIRD when implemented in isolation. For the second participant, RIRD was effective in decreasing vocal stereotypy and, therefore, RIRD was not combined with tact training. Given these findings, the effects of verbal operant training and RIRD on vocal stereotypy warrant additional research and suggest that RIRD would be effective in conjunction with EIBI during a variety of activities, including matching pictures, receptive identification, tacting, and other academic tasks, which were explored during the course of this study.
Chapter 2: Study Purpose

Numerous studies support the use of behavior analytic interventions to decrease aberrant vocal behavior displayed by individuals with ASD, including RIRD. Moreover, EIBI is an empirically-validated method for increasing appropriate language, social, and academic skills. Although it is common clinical practice to combine behavior-reduction techniques such as RIRD within skill acquisition programs, most published studies do not include both procedures. Additionally, while Ahearn et al. (2007) noted increases in language skills during the implementation of RIRD, frequency counts were used to measure gains. Currently, no studies have measured language gains following successful intervention for vocal stereotypy with standardized language measures. Colon et al. (2012) suggested that standardized developmental measures be completed with participants to better understand participant repertoires in order to predict how intervention may impact vocal stereotypy. Furthermore, Rogers and Vismara (2008) recommended that interventions demonstrate increases in both communicative behaviors and standard scores on developmental or language testing to increase the likelihood that children with ASD will achieve long-term success. Therefore, the goals of this study were (1) confirm the efficacy of RIRD during EIBI in decreasing vocal stereotypy and (2) to integrate the aforementioned gaps in the literature with respect to RIRD by evaluating the effects of this procedure on language skills, as measured by a standardized language tool.

Hypotheses

Specific Aim I was to determine if RIRD during EIBI would decrease vocal stereotypy to clinically acceptable levels in children with ASD. Given that multiple studies have demonstrated the effectiveness of RIRD in decreasing vocal stereotypy levels in children with ASD (e.g., Ahearn et al., 2007; Ahrens et al., 2011; Cassella et al., 2011; Colon
et al., 2012; Liu-Gitz & Banda, 2010; Miguel et al., 2009), it was hypothesized that children
in the experimental group would engage in lower levels of vocal stereotypy during EIBI
when RIRD was in effect than in baseline conditions.

Specific Aim II was to examine receptive and expressive language skills with
standardized language testing with the Preschool Language Scale-4 (PLS-4) to measure the
effects of RIRD. This aim was exploratory in nature because standardized language testing
had not been previously incorporated into RIRD studies. However, Ahearn et al. (2007) and
Miguel et al. (2009) observed concomitant increases in use of appropriate language when
participants received RIRD, as calculated by frequency counts during sessions. In the current
study, it was expected that higher standard scores would be observed at post-testing on
receptive and expressive subscales of the PLS-4 in the children with ASD who received
RIRD in conjunction with EIBI compared to participants in the clinical control group. This
hypothesis was developed based on data regarding how problems with attending during EIBI
tasks interfere with the learning process and children with problems related to attending
require additional procedures to acquire new language skills (Kodak et al., 2011). It is
therefore plausible that RIRD could lead to increases in attending because termination of
RIRD would occur only when the participants engaged in three correct responses without
vocal stereotypy. Because vocal stereotypy levels were expected to persist over time for the
participants within the clinical control group, problems with attending during EIBI programs
would also continue, likely resulting in no improvements on the PLS-4 for these participants
across administrations.

Specific Aim III was to identify whether there was a relationship between language
skills on the effectiveness of RIRD on suppression of vocal stereotypy. This aim was also
exploratory in nature. It was hypothesized that scores on receptive and expressive language subscales on the PLS-4 would be inversely related to frequency of vocal stereotypy. Moreover, it was hypothesized that the participant’s level of developmental functioning, as measured by the Early Learning Composite on the MSEL, would be positively correlated with language skills and negatively correlated with percentage of session with vocal stereotypy.
Chapter 3: Method

Participant Recruitment and Setting

Six children with ASD were recruited from the Autism Collaborative Center at Eastern Michigan University in Ypsilanti, Michigan, and the University of Nebraska Medical Center’s Munroe-Meyer Institute in Omaha, Nebraska. Participant inclusion criteria were (1) a chronological age between 36 to 68 months at the start of the study, (2) presence of vocal stereotypy that occurred at least several times per minute, and (3) a formal diagnosis of an ASD as determined by DSM-IV-TR criteria and standardized testing. Participants were assigned to one of three groups: the experimental group, who received RIRD to treat vocal stereotypy during their EIBI programming; the clinical control group; or the traditional treatment control group. Participants in the clinical control group received EIBI as part of their educational programming at their respective treatment centers. The purpose of the traditional treatment group was to determine maturational effects that occurred in language development during this age in children with ASD and no history of participating in an EIBI program.

The experimental design was quasi-experimental, in which participants were not randomly assigned but were instead participant or researcher-selected for groups (Shadish, Cook, & Campbell, 2002). Specifically, group assignment was based on current therapy enrollment (i.e., EIBI versus alternative therapies), and decisions regarding assignment to the experimental or clinical control group were also based on whether a participant had a pre-existing therapy goal related to decreasing vocal stereotypy. Participants in all groups were matched by age, sex, and level of developmental functioning as much as possible. Children with untreated severe behavior problems (e.g., self-injurious behavior or aggression that
caused visible bodily harm) were excluded from participation. All assessment and treatment sessions were conducted in the child’s treatment or educational setting with permission of participating agencies.

**Participants**

**Experimental Group.** “Steve” was 4 years and 5 months old at the start of the study. Steve frequently engaged in two- to three-word mands for preferred edibles and toys. Steve did not engage in many intraverbal skills at the beginning of the study. He attended an EIBI program for 6 hours per week, a combined speech-language, music therapy, and occupational group therapy program delivered once a week for 90 minutes, and a self-contained special education classroom for 20 hours per week. Table 1 shows a summary of participant characteristics, including age, group assigned, weekly hours of EIBI, current psychotropic medications, and tasks delivered during EIBI during sessions.

“Dylan” was 5 years and 7 months old at the start of the study. Dylan attended an EIBI program for 90 minutes per visit for four days per week for a total of 6 hours. Dylan was able to spontaneously mand with simple sentences (e.g., “I want goldfish.”) but did not spontaneously engage in tact behavior. Dylan readily engaged in echoic behavior and demonstrated emerging intraverbal skills, such as answering questions related to his name and age and filling in the last word of simple children’s songs. Dylan was in a self-contained special education classroom for 25 hours per week. In addition, he received ABA treatment for problems related to food selectivity at a clinic for pediatric feeding disorders for 1.5 hours per visit, three days per week.

**Clinical Control Group.** “George” was 5 years 0 months old at the start of the study. George engaged in spontaneous mands with up to four words (e.g., “I want jump,
please.”) and consistently repeated up to five words with echoic prompts. George engaged in tact behavior only when prompted but did not use tact frames, such as “I see ____” or “There’s a ____.” He was enrolled in an EIBI program for 6 hours per week (90-minute appointments for four days per week). George received school-based special education services in a self-contained classroom for children with ASD for 16 hours per week.

“Braxton” was 5 years and 7 months old when he began his participation, at which time he attended an EIBI program twice per week for 3 hours per visit, as well as a special education program for children with ASD for 20 hours per week. Braxton frequently engaged in mands with full sentences to request items and information from others. He also used full sentences to spontaneously tact items in his immediate environment. Prior to enrollment in present study, Braxton received intensive behavioral treatment for a pediatric feeding disorder and escape-maintained aggressive behavior (i.e., hitting, biting, kicking, and throwing objects towards others).

**Traditional Treatment Control Group.** “Adam” was 5 years 0 months of age at the start of participation. Adam’s spontaneous vocal language consisted primarily of two- to three-word mands (e.g., “all done”) and basic intraverbal skills. Adam attended a special education preschool program for 5 hours per day, four days per week (20 hours total). Additionally, he was enrolled in a weekly play group for preschoolers with ASD for 90 minutes per week, co-led by a music therapist, occupational therapist, and a speech-language pathologist.

“Noah” was 4 years and 3 months old at the time that the pre-testing measures were administered. Noah was able to mand in short sentences for items as well as information.
Table 1

*Summary of Participant Characteristics*

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Group</th>
<th>EIBI Hours Per Week</th>
<th>Medications</th>
<th>EIBI Tasks during Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>4:5</td>
<td>Experimental</td>
<td>6</td>
<td>None</td>
<td>Intraverbals (personal information; functions of items)</td>
</tr>
<tr>
<td>Dylan</td>
<td>5:7</td>
<td>Experimental</td>
<td>6</td>
<td>None</td>
<td>Tacts of pictures of common items</td>
</tr>
<tr>
<td>George</td>
<td>5:0</td>
<td>Clinical Control</td>
<td>6</td>
<td>None</td>
<td>Receptive identification of pictures of peers; matching pictures to objects</td>
</tr>
<tr>
<td>Braxton</td>
<td>5:7</td>
<td>Clinical Control</td>
<td>6</td>
<td>None</td>
<td>Intraverbals (personal information; object associations)</td>
</tr>
<tr>
<td>Adam</td>
<td>5:0</td>
<td>Traditional Control</td>
<td>0</td>
<td>None</td>
<td>N/A</td>
</tr>
<tr>
<td>Noah</td>
<td>4:3</td>
<td>Traditional Control</td>
<td>0</td>
<td>methylphenidate hydrochloride (18 mg)</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Noah was prescribed 18 milligrams of methylphenidate hydrochloride daily for reported problems related to hyperactivity and inattention and was monitored by a developmental-behavioral pediatrician. He was enrolled in a special education preschool program for 5 hours per day, four days per week and attended the same therapeutic play group as Adam and Steve for 90 minutes once per week.

**Assessment Measures**

All assessments were administered by advanced clinical psychology graduate students experienced in psychodiagnostic test administration and interpretation. All graduate students were supervised by a licensed psychologist. Figure 1 provides a diagram of the progression of study procedures for participants across the three groups.

**Pre-testing measures. Mullen Scales of Early Learning (MSEL; Mullen, 1995).**

At the start of the study, each participant was administered the MSEL, a standardized comprehensive measure of developmental functioning, to determine current overall level of functioning. This measure was selected because it was commonly used at the recruitment sites and it has a lower basal level than other measures of cognitive functioning normed for preschool-aged children, such as the Wechsler Preschool and Primary Scales of Intelligence (Wechsler, 2003). The MSEL is normed for children between the ages of birth to 68 months and is composed of four domains (*Visual Reception, Fine Motor, Receptive Language*, and *Expressive Language*) and an Early Learning Composite. With respect to psychometrics, the MSEL has sound internal consistency ($r = .91$) and inter-rater reliability ($r = .91$ to $.99$). Test-retest reliability is also satisfactory (mean stability coefficients = .76 for children 25 months and older). All scales are measured in T-scores, with a mean of 50 and standard deviation of 10. The Early Learning Composite has a mean of 100 and a standard deviation
of 15; this composite was used to determine each participant’s level of developmental functioning across all groups and was used in statistical analyses to evaluate the second hypothesis outlined in Specific Aim III.

**Autism Diagnostic Observation Schedule** (ADOS; Lord, Rutter, DiLavore, & Risi, 2002). To confirm a diagnosis of an ASD, the ADOS Module 1 or 2 was administered to all participants. Modules were selected based on the child’s estimated level of spontaneous language. Module 1 is designed for children who do not have vocal speech or communicate primarily with single words, whereas Module 2 is for children who vocalize with phrases but do not speak flexibly in full sentences. George and Dylan were administered Module 1 and Braxton, Steve, Adam, and Noah received Module 2. The ADOS was selected for use in this study to confirm diagnoses for participants because it is considered to be a “gold-standard,” play-based, standardized assessment of social, communication, and play skills. This measure is composed of three domains: Social Interaction, Communication, and Stereotyped Behaviors/Circumscribed Interests. The format of the ADOS consists of several standardized activities designed to occasion joint attention, communicative gestures, requesting behavior, and functional and imaginative play. Inclusion of assessment activities varies depending on the child’s level of expressive language and consists of activities such as a free-play component, pretend birthday party, snack, puzzle game, bubbles, balloons, and remote-controlled toys. The ADOS has been demonstrated to have solid psychometric properties, with consistently high interrater reliability across Modules 1 through 4 (mean exact agreement = 88.2 to 91.5%). Test-retest correlations yield superior stability in the Social Interaction and Communication domains ($r = 0.78$ to 0.73, respectively) and good stability in the Stereotyped Behaviors domain ($r = 0.59$). After the ADOS was administered,
the revised algorithm (Gotham, Risi, Pickles, & Lord, 2007) was used to compute scores across Communication, Reciprocal Social Interaction, and Repetitive/Restrictive Behavior domains as well as a combined total score. Results of the ADOS indicated that all participants received a combined total score above the clinical cut-off suggestive for autism.

**Pre- and Post-Testing.** *Preschool Language Scale- Fourth Edition* (PLS-4; Zimmerman, Steiner, & Pond, 2002). The PLS-4 was administered to all participants and was the only assessment measure that was repeated at post-testing (approximately five months later). This measure was selected to evaluate participants’ receptive and expressive language skills because it has been demonstrated to be suitable for use with preschool-aged children with ASD across a large sample (Volden et al., 2011). The PLS-4 is normed for children between birth to 6 years 11 months. The Auditory Comprehension and Expressive Communication domains are the two subscales that compose the PLS-4, which are used to calculate the Total Language Score. Test-retest reliability for the PLS-4 has been demonstrated to be stable, with stability coefficients ranging between .82 to .95 for subscales and .90 to .97 for the Total Language Score. In addition, internal consistency, as calculated by split-half reliability, has been shown to be solid across subscales and ages of examinees ($r = .83$ to $.95$ for Auditory Comprehension; $r = .82$ to $.95$ for Expressive Communication) and the Total Language Score ($r = .91$ to .97). Agreement across independent raters was 99%, indicating a high level of inter-rater reliability. Raw scores for each scale were calculated into standard scores with a mean of 100 and standard deviation of 15, and these standard scores were used as the dependent variable. Standard scores ranging from 85 to 115 were considered to be within normal limits.
Procedures for Experimental and Clinical Control Groups

**Functional Analysis.** To determine the function of vocal stereotypy for each participant in the experimental and clinical control groups, a functional analysis consisting of *attention, demand, ignore/no-consequence*, and *play* conditions, as described earlier, was conducted (Iwata et al., 1982/1994). Conditions were presented in a multielement design and all sessions were 5 minutes in duration. The condition order was randomized and determined prior to implementation. Functional analysis sessions were conducted until a stable pattern of responding was observed or there were consistently high levels of vocal stereotypy across extended *ignore or no-consequence* sessions.

**Attention.** At the beginning of the session, the therapist sat in close proximity to the participant on a couch or at a table ("close" means approximately one adult arm’s length). The child was provided with two toys that were identified by the parent as being moderately preferred. The therapist engaged in 1-minute of high-quality attention with the participant and then stated, “I have some work to do; you can play with your toys quietly,” turned his or her head away from the child, and pretended to read a magazine. Contingent on the occurrence of vocal stereotypy, the therapist directed his or her attention to the child and delivered verbal reprimands such as, “Please be quiet, I am busy doing my work” and “I don’t like it when you make those noises.” Reprimands were continued for 20 seconds, at which time the therapist turned away from the participant again, and resumed reading a magazine. This procedure was repeated contingent on further displays of vocal stereotypy. All other inappropriate behaviors (e.g., body rocking, throwing toys) were ignored.

**Demand.** The purpose of this condition was to rule out escape-maintained vocal stereotypy. The therapist sat across from the child at a preschool-sized table with academic
materials. Mastered teaching tasks were selected based on the child’s current EIBI therapy curriculum, and included activities such as gross motor imitation, visual performance (e.g., puzzles, block design, sorting shapes into a container), receptive commands (e.g., “Stand up,” “Turn around,” etc.), and sorting items by function or class. The therapist used a three-step prompting procedure during each trial (Tucker & Berry, 1980). Specifically, the therapist asked the participant to complete the task (e.g., “Put the puzzle together”). If he did not comply within 5 seconds, the therapist repeated the vocal prompt while modeling how to complete the task. If, after an additional 5 seconds, the participant did not initiate the appropriate response, the therapist used physical guidance to assist the child with completing the response by placing his or her hand over the child’s hand and engaging in the task. In the event that the child engaged in vocal stereotypy during any of these prompts, the therapist provided him with a 20-second break. The break was signified by the therapist saying, “Okay, you do not have to,” removing the task materials from the participant’s view, and turning his or her back to him for the 20-second allotted break. At the conclusion of the break, a new task was initiated with the three-step prompting hierarchy. Descriptive verbal praise was delivered when the child completed the correct response (e.g., putting the puzzle together) without physical guidance. Examples of verbal praise included, “Nice job putting the puzzle together!” and “That’s right, that is the word dog!” Tasks requiring a vocal response were not included in this test condition because the therapist was not able to use physical prompting. The learning tasks were initiated as often as allowed during the 5-minute session. No contingencies were in place for other inappropriate behaviors (e.g., out-of-seat behavior, aggression, motor stereotypy).
**Ignore.** An *ignore* condition was implemented with Dylan, Henry, and Braxton. During this condition, the participant and therapist were in a barren room, in which no toys, furniture, or other objects were present. The therapist did not interact with the participant in any manner. No programmed contingencies were in place for vocal stereotypy. All appropriate requests for attention or to escape from the room were ignored, as well as any inappropriate behavior. If a participant attempted to exit the therapy room before the session was over, a data collector behind a one-way observation mirror locked the door until the child released his hand from the door handle.

**No-consequence.** For Steve, a *no-consequence* condition (Athens et al., 2008) was conducted to determine if his vocal stereotypy was maintained by automatic reinforcement. This condition was conducted instead of a traditional *ignore* condition because Steve did not have a history of being in an impoverished natural environments and the clinic where he received his treatment did not an appropriate room where all materials could be removed. During this condition, Steve was in a play room with access to five toys. The therapist was present but did not interact with Steve in any manner. No programmed contingencies were in place for vocal stereotypy. Mands and other appropriate vocalizations were also ignored.

**Play.** This condition served as the control procedure. The therapist was present in the room and maintained close proximity (i.e., no more than two feet away) to the child. The child had continuous access to a variety of preferred toys, which were selected according to parental report or multiple-stimulus without replacement preference assessments (DeLeon & Iwata, 1996) that were conducted during the child’s typical EIBI programming. During *play* sessions, the therapist did not deliver educational tasks or vocal prompts and noncontingent attention in the form of verbal praise was delivered at least once every 30 seconds (e.g., “I
love the way that you are playing with your blocks”). The participant was permitted to move freely around the room. No programmed contingencies were in place for vocal stereotypy.

**Experimental Design.** A withdrawal design was implemented across participants in the experimental group to analyze the effects of RIRD on vocal stereotypy during EIBI. For the clinical control group, George and Braxton experienced an extended baseline condition during their typical EIBI sessions at their treatment center.

**Baseline during EIBI.** Baseline sessions were conducted during the participant’s EIBI appointments, at which time the participant was required to perform tasks outlined by his educational programming, such as matching, receptive identification, tact training, and intraverbal training of personal information (e.g., “How old are you?”). All sessions were 5 minutes in duration. Any occurrences of vocal stereotypy were ignored by the therapist and recorded on a data sheet (for Steve only; please see Appendix D) or with a computerized data collection system by an observer other than the therapist.

**EIBI + Vocal RIRD.** RIRD was implemented based on the procedures outlined by Ahearn et al. (2007). At the start of the session, EIBI tasks were initiated in a manner similar to the baseline condition described above. Reinforcement and prompting procedures used were consistent with those specified by the participant’s EIBI curriculum (e.g., reinforcement delivered on a fixed-ratio 1 schedule for correct, unprompted responses). Contingent on the occurrence of vocal stereotypy, the therapist suspended EIBI tasks and began initiating RIRD trials in the form of vocal imitation tasks; specifically, in a neutral tone of voice, the experimenter secured the child’s attention while initiating eye contact and provided echoic prompts (e.g., “Say baby”). Vocal imitation tasks were selected a priori, based on the child’s current language skills and mastered words according to his existing EIBI curriculum. RIRD
tasks were delivered until the participant engaged in correct imitation of three words in the absence of vocal stereotypy. If the participant did not say the word correctly following the echoic prompt (i.e., the child did not say anything, said a different word, or engaged in vocal stereotypy), the therapist repeated the same word until he echoed the word without vocal stereotypy. Once the participant met the discontinuation criteria for RIRD, EIBI tasks were resumed. For instance, if the child was given the discriminative stimulus, “What is it?” while the therapist held up a picture of a cup and then the child engaged in vocal stereotypy, the therapist removed the picture from the child’s view and conducted RIRD trials until he echoed three words correctly without further display of vocal stereotypy; contingent on meeting this criteria, the therapist would re-present the picture of the cup and repeat the trial. If the participant engaged in vocal stereotypy during the inter-trial interval, the therapist implemented RIRD tasks immediately. Verbal praise and edible reinforcement were delivered contingently on correct responding during EIBI tasks only. All sessions were 5 minutes in duration. Sessions were conducted one to two times per week, with sessions not exceeding 30 minutes per day.

**EIBI + Motor RIRD.** Dylan also was exposed to RIRD tasks requiring gross motor imitation. Similar to the EIBI + Vocal RIRD condition, EIBI tasks were initiated at the beginning of the session; however, contingent on vocal stereotypy, gross motor imitation tasks were delivered until Dylan met the discontinuation criteria described above (i.e., correct responding for three consecutive tasks without emitting vocal stereotypy). Gross motor imitation tasks consisted of the therapist delivering the instruction, “Do this,” while demonstrating a motor action, such as knocking on the table or putting her hand on top of her head. Receptive commands (e.g., “Clap your hands,” “Touch your head,” “Jump,” etc.) were
excluded from this condition because Dylan was currently learning such tasks during his EIBI programming and the experimenter wanted to increase the probability that Dylan would discriminate between EIBI trials and RIRD tasks for vocal stereotypy. A three-step prompting procedure was utilized (Tucker & Barry, 1980), similar to the prompts used during the demand sessions of the functional analysis, in which the therapist progressed through a verbal, gestural, and then physical prompting sequence if Dylan did not comply with the instruction within 5 seconds. Verbal praise and edible reinforcement were only delivered contingent on correct responding during EIBI tasks; that is, verbal praise was not provided when Dylan correctly imitated actions without physical guidance during RIRD trials.

Figure 1. Outline of study procedures across groups.
Data Collection and Interobserver Agreement

Vocal stereotypy was defined as the occurrence of any repetitive or noncontextual speech. Immediate echolalia, the immediate repetition of a word or phrases (McMorrow, Foxx, Faw, & Brittle, 1987), was excluded from this definition because it was often unclear if the participant was emitting vocal stereotypy or engaging in functional speech. For instance, following the instruction, “Touch your nose,” Dylan often would say “nose” while pointing to his nose; in this case, it was unclear if Dylan was engaging in vocal stereotypy or tactualing his nose, which would be an appropriate vocalization. All participants engaged in vocal stereotypy in the forms of delayed echolalia, nonfunctional noises, and laughing in the absence of a humorous event. Common examples of palilalia included reciting lines from children’s television programs and movies and repeating sentences heard from caregivers, such as listing the characters from Thomas the Tank Engine® (e.g., “Thomas and Toby and Henry and Gordon and Percy and Harold…”), reciting phrases from the movie Toy Story® (e.g., “To infinity and beyond!”), and echoing caregiver reprimands (e.g., “Shut up. Don’t say shut up, shut up’s a bad word!” and “No hitting. Do you want to go to time-out?”).

Vocal stereotypy data were collected using a 10-second partial-interval recording (PIR) procedure. PIR is a time-sampling method, in which observers denote if the target response occurs at any time during a specified interval (Cooper et al., 2007). Data were then reported as a percentage of total intervals in which vocal stereotypy was observed. For Dylan, George, and Braxton, vocal stereotypy data were collected on laptop computers with a computer program designed to track frequency and duration data (DataPal). During computerized data collection, observers began collecting data at the immediate onset of vocal stereotypy and turned off the duration key if vocal stereotypy did not occur for 1 second.
Steve’s data were collected using paper and pencil procedures, in which observers recorded the occurrence of vocal stereotypy by writing a “V” in the box that denoted the specific 10-second interval in which Steve engaged in vocal stereotypy. If Steve did not engage in vocal stereotypy during an interval, data collectors indicated the nonoccurrence by writing a “—” in the respective interval. Please refer to Appendix D for an example of this data sheet.

To measure interobserver agreement, two observers independently recorded at least 33% of all functional analysis, baseline, and treatment sessions. Interobserver reliability was calculated by dividing the number of intervals with agreements by the total number of intervals with agreements plus disagreements and multiplying by 100%. For Steve, mean interobserver agreement was 92.5% across functional analysis conditions (range = 85% to 100%), 97% across baseline sessions (range = 90% to 100%), and 97.5% across EIBI + RIRD sessions (range = 95% to 100%). With regard to Dylan’s sessions, mean interobserver agreement was 88% across functional analysis conditions (range = 81% to 94%), 91% across baseline sessions (range = 86% to 95%), and 85.5% across EIBI + RIRD sessions (range = 82.5% to 88%). For George, mean interobserver agreement was 88% across functional analysis conditions (range = 83% to 92%) and 96% across extended baseline sessions during EIBI (range = 78% to 100%). Lastly, for Braxton’s sessions, mean interobserver agreement was 92% across functional analysis conditions (range = 89% to 97%) and 89% across extended baseline during EIBI condition (range = 82% to 94.5%).

Data Analysis

Specific Aim I. All functional analysis, baseline during EIBI, and EIBI + RIRD sessions were graphed and analyzed with visual inspection procedures, as described by Parsonson and Baer (1978, 1992).
Specific Aim II. To assess for differences in standard scores achieved on the first and second administrations of the PLS-4 across the three groups, effect sizes were calculated with Cohen’s $d$ (Cohen, 1988, 1992, 1994). Although Wilcoxon signed-rank tests were originally proposed to evaluate differences in PLS-4 language scales from pre- to post-testing across the experimental and clinical control groups, such statistical analyses could not be calculated because less than 6 participants were in each group (Howell, 2007).

Specific Aim III. Pearson’s $r$ correlation coefficients were calculated to identify the relationships between (1) vocal stereotypy level and language skills, (2) vocal stereotypy level and developmental level of functioning, and (3) developmental level of functioning and language skills. For these analyses, vocal stereotypy level was calculated by determining the mean percentage of intervals of vocal stereotypy across the last 5 sessions for each participant within the experimental and clinical control groups; these sessions were selected because they were conducted within two weeks prior to post-testing. Language skills were measured by the PLS-4 Auditory Comprehension, Expressive Communication, and Total Language standard scores and developmental level of functioning was measured by the Early Learning Composite on the MSEL.
Chapter 4: Results

Specific Aim I

The results of the functional analysis for the participants in the experimental group are depicted in Figure 2. Vocal stereotypy patterns were undifferentiated across functional analysis conditions for both participants. For Steve (shown in the top panel of Figure 2), vocal stereotypy occurred at moderate levels in the demand condition, moderate to high levels in the attention condition, and high levels in the ignore condition. An increasing trend was also observed during the play condition. This pattern warranted the implementation of an extended no-consequence condition series; results of this condition demonstrated that Steve continued to engage in consistently high levels of vocal stereotypy across sessions \(M = 91.67\% \text{ of intervals}\), supporting the hypothesis that his vocal stereotypy was maintained by automatic reinforcement. During Dylan’s functional analysis, moderate levels of vocal stereotypy were observed during the attention, demand, and play conditions. Dylan engaged in moderate to high levels of vocal stereotypy in the ignore condition of the functional analysis and levels of vocal stereotypy remained high during the extended ignore condition phase \(M = 66.67\% \text{ of intervals}\), indicating that his vocal stereotypy also most likely served an automatic function.

Figure 3 shows the functional analysis results for George and Braxton, the participants in the clinical control group. Similar to the participants in the experimental group, George engaged in high levels of vocal stereotypy across all conditions of the functional analysis. He also continued to engage in high levels of vocal stereotypy in the
Figure 2. Functional analysis results for participants in the experimental group.
Figure 3. Functional analysis results for participants in the clinical control group.
extended *ignore* phase ($M = 80\%$ of intervals). These data revealed that his vocal stereotypy was most likely automatically-maintained. As shown in the bottom panel, Braxton engaged in lower levels of vocal stereotypy across all conditions compared to George. Braxton engaged in low levels of vocal stereotypy across the *attention* and *play* conditions during his functional analysis, with an increasing trend noted during the final two *demand* sessions. Variable responding was observed during *ignore* sessions ($M = 44.17\%$ of intervals) with low to moderate levels across sessions. To confirm that vocal stereotypy persisted in the absence of social consequences, an extended *ignore* series was introduced and vocal stereotypy levels continued to persist, with an increasing trend observed ($M = 30\%$ of intervals).

The results of the EIBI + RIRD conditions are shown in Figure 4. Steve (top panel) engaged in moderate levels of vocal stereotypy during the initial baseline during EIBI phase; these levels were somewhat lower than those observed during the extended *ignore* series of the functional analysis (i.e., $M = 60.5\%$ of intervals during baseline versus $M = 91.67\%$ of intervals during *ignore* sessions), suggesting that EIBI tasks may have had some decelerative effect on vocal stereotypy. When RIRD procedures were combined with EIBI, a 77\% reduction in vocal stereotypy was noted across sessions ($M = 13.8\%$ of intervals). During the reversal back to the baseline condition, vocal stereotypy levels immediately returned to moderate levels ($M = 48.3\%$ of intervals). A similar decrease in vocal stereotypy levels was observed when RIRD was reintroduced and stereotypy levels remained low for several consecutive sessions ($M = 16.7\%$ of intervals).

A different pattern of responding was noted with Dylan, as shown in the bottom panel of Figure 4. Dylan engaged in moderate levels of vocal stereotypy in the baseline condition
when EIBI tasks were delivered, with an increasing trend noted during the final sessions ($M = 47.4\%$ of intervals). However, the introduction of vocal RIRD to EIBI sessions did not lead to reductions in Dylan’s vocal stereotypy levels ($M = 47.9\%$ of intervals). Likewise, when RIRD tasks were switched to motor imitation tasks during EIBI, Dylan continued to engage in similar levels of vocal stereotypy compared to baseline ($M = 47.5\%$ of intervals).

Figure 5 depicts the results of the extended baseline during EIBI condition for the participants in the clinical control group. As shown in the top panel, George engaged in moderate to high levels of vocal stereotypy across all 26 sessions ($M = 70\%$ of intervals). These data indicate that vocal stereotypic responses maintained over time and did not decrease. Braxton’s data are denoted in the bottom panel. Braxton engaged in variable levels of vocal stereotypy during EIBI ($M = 30.96\%$ of intervals; range $= 10\%$ to $75\%$ of intervals).

**Specific Aim II**

Tables 2, 3, and 4 summarize the pre- and post-testing results for the PLS-4 across all participants. It was hypothesized that higher standard scores would be observed at post-testing on the Auditory Comprehension and Expression Communication domains of the PLS-4 in participants who received EIBI + RIRD compared to those in the clinical control group. Effect sizes were computed using Cohen’s $d$ to determine if there were any changes across pre- and post-testing standard scores on the PLS-4. Large effects were found in the experimental group and very large effects were noted in the traditional treatment control group on the Auditory Comprehension, Expressive Communication, and Total Language standard scores from pre- to post-testing. For the control group, small effects
Figure 4. Results of the baseline during EIBI and EIBI plus RIRD conditions for participants in the experimental group.
Figure 5. Results of the extended baseline condition during EIBI for participants in the clinical control group.
Table 2

PLS-4 Results across Participants: Auditory Comprehension

<table>
<thead>
<tr>
<th>Name</th>
<th>Group</th>
<th>AC Time 1</th>
<th>AC Time 2</th>
<th>Change</th>
<th>Effect Size d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>Experimental</td>
<td>74</td>
<td>84</td>
<td>+10</td>
<td>0.70</td>
</tr>
<tr>
<td>Dylan</td>
<td>Experimental</td>
<td>50</td>
<td>50</td>
<td>+0</td>
<td></td>
</tr>
<tr>
<td>George</td>
<td>Clinical Control</td>
<td>76</td>
<td>66</td>
<td>-10</td>
<td>0.20</td>
</tr>
<tr>
<td>Braxton</td>
<td>Clinical Control</td>
<td>67</td>
<td>85</td>
<td>+18</td>
<td></td>
</tr>
<tr>
<td>Adam</td>
<td>Traditional Control</td>
<td>68</td>
<td>91</td>
<td>+23</td>
<td>10.14</td>
</tr>
<tr>
<td>Noah</td>
<td>Traditional Control</td>
<td>81</td>
<td>101</td>
<td>+20</td>
<td></td>
</tr>
</tbody>
</table>

Note: AC = Auditory Comprehension scale ($M = 100; SD = 15$); Time 2 occurred approximately five months later. Effect size is calculated as Cohen’s $d$. 
Table 3

**PLS-4 Results across Participants: Expressive Communication**

<table>
<thead>
<tr>
<th>Name</th>
<th>Group</th>
<th>EC Time 1</th>
<th>EC Time 2</th>
<th>Change</th>
<th>Effect Size $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>Experimental</td>
<td>73</td>
<td>85</td>
<td>+12</td>
<td>-0.71</td>
</tr>
<tr>
<td>Dylan</td>
<td>Experimental</td>
<td>50</td>
<td>50</td>
<td>+2</td>
<td></td>
</tr>
<tr>
<td>George</td>
<td>Clinical Control</td>
<td>57</td>
<td>50</td>
<td>-7</td>
<td>-0.19</td>
</tr>
<tr>
<td>Braxton</td>
<td>Clinical Control</td>
<td>65</td>
<td>77</td>
<td>+12</td>
<td></td>
</tr>
<tr>
<td>Adam</td>
<td>Traditional Control</td>
<td>54</td>
<td>62</td>
<td>+8</td>
<td>-0.91</td>
</tr>
<tr>
<td>Noah</td>
<td>Traditional Control</td>
<td>82</td>
<td>83</td>
<td>+1</td>
<td></td>
</tr>
</tbody>
</table>

*Note: EC = Expressive Communication scale ($M = 100; SD = 15$); Time 2 occurred approximately five months later. Effect size is calculated as Cohen’s $d$.  

Table 4

*PLS-4 Results across Participants: Total Language Score*

<table>
<thead>
<tr>
<th>Name</th>
<th>Group</th>
<th>Total Lang. Time 1</th>
<th>Total Lang. Time 2</th>
<th>Change</th>
<th>Effect Size d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>Experimental</td>
<td>71</td>
<td>94</td>
<td>+23</td>
<td>0.71</td>
</tr>
<tr>
<td>Dylan</td>
<td>Experimental</td>
<td>50</td>
<td>50</td>
<td>+0</td>
<td></td>
</tr>
<tr>
<td>George</td>
<td>Clinical Control</td>
<td>63</td>
<td>54</td>
<td>-9</td>
<td>-0.22</td>
</tr>
<tr>
<td>Braxton</td>
<td>Clinical Control</td>
<td>62</td>
<td>79</td>
<td>+17</td>
<td></td>
</tr>
<tr>
<td>Adam</td>
<td>Traditional Control</td>
<td>57</td>
<td>80</td>
<td>+23</td>
<td>-1.21</td>
</tr>
<tr>
<td>Noah</td>
<td>Traditional Control</td>
<td>85</td>
<td>91</td>
<td>+6</td>
<td></td>
</tr>
</tbody>
</table>

*Note: M = 100; SD = 15; Time 2 occurred approximately five months later. Effect size is calculated as Cohen’s d.*
Table 5

*MSEL Results across Participants*

<table>
<thead>
<tr>
<th>Name</th>
<th>Group</th>
<th>Visual Reception&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Fine Motor&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Receptive Language&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Expressive Language&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Early Learning Composite&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>Experimental</td>
<td>34</td>
<td>20</td>
<td>20</td>
<td>32</td>
<td>60</td>
</tr>
<tr>
<td>Dylan</td>
<td>Experimental</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>George</td>
<td>Clinical Control</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>Braxton</td>
<td>Clinical Control</td>
<td>40</td>
<td>45</td>
<td>33</td>
<td>20</td>
<td>71</td>
</tr>
<tr>
<td>Adam</td>
<td>Traditional Control</td>
<td>20</td>
<td>28</td>
<td>29</td>
<td>31</td>
<td>51</td>
</tr>
<tr>
<td>Noah</td>
<td>Traditional Control</td>
<td>30</td>
<td>23</td>
<td>38</td>
<td>32</td>
<td>65</td>
</tr>
</tbody>
</table>

*Note:* <sup>a</sup> For all MSEL subscales, $M = 50; SD = 10$; <sup>b</sup> $M = 100; SD = 15$
were observed across all scales on the PLS-4.

**Specific Aim III**

To assess whether there was a relationship between vocal stereotypy level and language scores on the PLS-4 at post-testing, Pearson correlations were conducted. It was hypothesized that standard scores on receptive and expressive language subscales of the PLS-4 would be inversely related to frequency of vocal stereotypy at post-testing for participants within the experimental and clinical control groups. In addition, it was hypothesized that participant developmental level of functioning, as measured by the Early Learning Composite on the MSEL, would be positively related to language skills and negatively related to percentage of intervals with vocal stereotypy. Results for the MSEL for all participants are shown in Table 5. Table 6 depicts the results of the Pearson correlation analysis across vocal stereotypy level, PLS-4 post-testing language scales, and developmental level of functioning. No associations were found between vocal stereotypy level and language skills or vocal stereotypy level and developmental level. A strong positive correlation was found between expressive communication skills at post-testing and developmental level of functioning $r (6) = .837, p < .05$. 
Table 6

*Pearson Correlations between Vocal Stereotypy, Language Skills, and Developmental Level of Functioning*

<table>
<thead>
<tr>
<th></th>
<th>Vocal Stereotypy</th>
<th>Developmental Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Comprehension Skills</td>
<td>-0.620</td>
<td>0.628</td>
</tr>
<tr>
<td>Expressive Communication Skills</td>
<td>-0.897</td>
<td>0.837*</td>
</tr>
<tr>
<td>Total Language Score</td>
<td>-0.894</td>
<td>0.673</td>
</tr>
<tr>
<td>Developmental Level</td>
<td>-0.578</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Note: *p < .05
Chapter 5: Discussion

Vocal stereotypy is a problematic behavior commonly displayed in children with ASD, which has been demonstrated to interfere with establishing positive social interactions and skill acquisition during different types of academic programming (Lanovaz & Sladeczek, 2012). This study sought to investigate the effects of RIRD on vocal stereotypy when implemented during EIBI programs with preschool-aged children with ASD. RIRD is an intervention with growing empirical evidence for reducing vocal stereotypy maintained by automatic reinforcement. For example, RIRD has been shown to be effective across clinical and educational settings (Liu-Gitz & Banda, 2010), with different topographies of both vocal and motor stereotypies (Ahrens et al., 2011), and when compared to pharmacological intervention (Miguel et al., 2009). Given that EIBI is the treatment of choice for children with ASD for teaching a variety of language, social, and academic skills, determining the effectiveness of interventions for vocal stereotypy during EIBI is important so that clinicians can minimize barriers to learning and maximize therapeutic outcomes.

Effectiveness of RIRD on Vocal Stereotypy during EIBI

The hypothesis that RIRD would be successful in decreasing vocal stereotypy to acceptable levels was partially supported. Specifically, RIRD was effective in diminishing vocal stereotypy levels for one of two participants within the experimental group. For Steve’s data, these results replicate the findings of Colon et al. (2012) in that EIBI + RIRD was successful in reducing vocal stereotypy to clinically acceptable levels. In addition, this study extended the results of Colon et al. by incorporating intraverbal training during Steve’s sessions. Moreover, verbal operant training was insufficient in significantly decreasing vocal
stereotypy levels for both Steve and Dylan, which was consistent with the findings for two of the three participants in the Colon et al. investigation.

Though it is unclear why RIRD was ineffective in reducing Dylan’s vocal stereotypy during EIBI, it is possible that the imitation tasks selected during RIRD trials did not function as punishers. Instead, vocal and motor imitation tasks were selected because Dylan had a strong imitative repertoire. Recent research by Torres-Viso, Sloman, and Shulman (2012) demonstrated that RIRD was ineffective for a girl with ASD when imitation tasks were used as the contingent demands. A negative reinforcement test was then conducted to identify aversive tasks and results indicated that intraverbal tasks requiring the participant to fill in the blanks for common children songs (e.g., the therapist said, “Twinkle, twinkle, little” and the child was required to say “star”) were non-preferred. When this task was incorporated into RIRD trials contingent on vocal stereotypy, clinically significant decreases were observed. In contrast, Liu-Gitz and Banda (2010) reported reductions in vocal stereotypy when RIRD was implemented with tasks related to answering questions regarding the participant’s preferred topics (e.g., “Who is your favorite character in Toy Story?”). The results of these studies suggest that the RIRD tasks selected are idiosyncratic to the child, similar to a preference assessment.

It is possible that Dylan’s tacting program implemented during EIBI sessions was equally or less preferred than the vocal and motor imitation tasks designated for RIRD, which could explain why a reduction in vocal stereotypy was not observed. Alternatively, Dylan received frequent access to reinforcers when RIRD was delivered during EIBI, because reinforcement was delivered for correct tact responses (i.e., independent and prompted responses), thereby lessening the aversive quality of RIRD for the overall session.
Within the EIBI + RIRD conditions, Dylan may have had problems with discrimination between RIRD and EIBI tasks, which also could account for the ineffectiveness of RIRD in decreasing his vocal stereotypy.

An alternative interpretation for Dylan’s results during EIBI is that his aberrant vocal responding may have been vocal tics rather than vocal stereotypy. Although Gilbert (2006) published some guiding principles for differentiating between these two behaviors, no systematic method has been developed to distinguish vocal stereotypy from vocal tics thus far. Behavior-analytic intervention studies on vocal stereotypy do not include criteria for discriminating between vocal stereotypy and vocal tics; instead, operational definitions for vocal stereotypy are based on topography, whether the utterance is irrelevant to the current context, and that the vocalization appears to be nonfunctional. During RIRD trials in the present study, it was particularly difficult to determine if Dylan’s vocal responding was rhythmic in nature because therapists interrupted the behavior immediately. However, vocal responding occurred at consistently high percentages in the ignore condition of the functional analysis; given the high level of responding during these sessions, Dylan’s vocal behavior was likely rhythmic, which is a distinctive feature of vocal stereotypy. Anecdotal observations also indicated that Dylan would sometimes engage in vocal stereotypy during functional speech, which is not a characteristic of a vocal tic (Gilbert, 2006). In addition, the onset for Dylan’s atypical vocal behaviors was prior to age three, which suggests that these behaviors were more likely vocal stereotypy (Freeman et al., 2010).

Habit reversal is a primary behavioral treatment package for vocal tics (Carr & Chong, 2005) that is comprised of multiple components, including a functional assessment, awareness training, and implementation of a competing response (Woods et al., 2008;
Woods, Twohig, Flessner, & Roloff, 2003). Habit reversal techniques require an individual to be able to self-monitor their tics, which is not a requirement for RIRD. However, it is unclear whether participants in the current study were aware of their vocal stereotypy prior to the implementation of RIRD and it is possible that RIRD contingencies served as a type of awareness training, teaching participants to suppress vocal stereotypy until sessions had elapsed.

In the ABA literature, tics are sometimes hypothesized as operant behavior (Richman & Lindauer, 2002; Watson & Sterling, 1998), similar to hypothesized mechanisms for stereotypy. Though tics are often conceptualized as involuntary behaviors (American Psychiatric Association, 2000), tics are at least “under partial voluntary control” because many children are able to suppress their tics temporarily (Swain et al., 2007). As previously noted, tics occur more likely to occur when individuals are in relaxed states, such as when they engage in solitary activities (American Psychiatric Association, 2000); if vocal tics are an operant behavior, this information suggests that tics may be automatically-reinforced for certain individuals, similar to vocal stereotypy. However, it is also possible to suppress and modify ordinarily involuntary, non-operant behavior, such as breathing rate, indicating that such criterion may also not clearly distinguish among the behavior types.

Similarities are also noted across treatments for vocal stereotypy and vocal tics. Some behavioral treatment studies of vocal tics have had mixed results of effectiveness (e.g., Woods & Twohig, 2002), indicating like vocal stereotypy, vocal tics are a challenging behavior to treat. Moreover, treatments for vocal stereotypy and tic suppression sometimes contain similar components, such as DRO. Woods and Himle (2004) applied DRO and verbal instruction to vocal tics in four children with Tourette’s Disorder (without ASD). In
the verbal instruction condition, participants were told to, “Do whatever you need to do to keep your tics from happening during the next 5 minutes,” whereas in the verbal instruction plus DRO condition, participants were given the same instruction and also were delivered tokens on a 10-second DRO schedule for the absence of tics. Results showed that the combined condition was more effective in suppressing tics across all participants. Likewise, Himle, Woods, and Bunaciu (2008) found that a DRO-based token economy was successful in reducing tics in three of four school-age children with Tourette’s Disorder, while tics occurred at a similar rate to baseline when tokens were delivered on a response-independent schedule.

**Effects of RIRD on Standardized Scores of Language Skills**

It was predicted that children who were exposed to RIRD during EIBI would achieve higher increases in standard scores from pre- to post-testing on the PLS-4 Auditory Comprehension and Expressive Communication domains as well as the Total Language score compared to children within the clinical control group. This hypothesis was only partially supported because participants in the traditional treatment control group outperformed all participants on the PLS-4 from Time 1 to Time 2. Across the three groups, one participant in each group demonstrated considerable increases on the PLS-4 Total Language score. Within the experimental group, Steve’s performance on the PLS-4 indicated that his Total Language score increased by 23 points between the first and second administrations, whereas no change was observed across Dylan’s pre- and post-testing standard scores on the PLS-4. However, a large effect size was still noted ($d = -0.71$). For participants in the clinical control group, George’s Total Language standard score declined by 9 points, in contrast to the 17-point increase noted in Braxton’s Total Language standard
score. Finally, a similar pattern was observed with participants in the traditional treatment control group, with a 23-point increase on the Total Language score for Adam and a smaller increase in Total Language score for Noah (6 points) from Time 1 to Time 2; particularly, large effects were noted in this group on the Auditory Comprehension scale. Given the small sample size of this study, it is difficult to draw conclusions regarding the effects of suppression of vocal stereotypy on language skills, as there was low power (Cohen, 1992).

The observed decreases in receptive and expressive language scores for George was unexpected given that he was receiving EIBI, which is considered an empirically-supported treatment for teaching a variety of language skills to children with ASD. However, as shown in Figure 5, George engaged in high levels of vocal stereotypy throughout his EIBI sessions, which may have impacted his ability to attend to relevant stimuli during teaching trials, affecting the learning process. Alternatively, it is also possible that George’s vocal stereotypy interfered with his performance during the PLS-4 and results obtained may have been an underestimate of George’s language abilities.

The significant increases in PLS-4 standards scores on the Auditory Comprehension ($d = 10.14$) and Total Language ($d = -1.20$) domains for Adam and Noah, the participants within the traditional treatment control group, was a noteworthy and surprising finding. Steve, Adam, and Noah attended a weekly 90-minute combined occupational, speech/language, and music therapy group, which may have led to improved performance on the second PLS-4 administration. In addition, increases in standard scores noted in Steve’s language PLS-4 profile may have possibly been due to therapeutic gains from his group therapy rather than the combination of EIBI + RIRD. Information regarding language
maturation in young children with ASD is unavailable and the results of this study indicate that additional data are needed to understand language trajectories in this population.

The PLS-4 was selected as the dependent measure because it has been demonstrated as a valid language assessment tool for preschoolers with ASD. Volden and colleagues (2011) conducted a large-scale study with 294 children between the ages of 2 to 5 who had been recently diagnosed with ASD. The purpose of this study was to verify the administration of the PLS-4 as an appropriate assessment for children with ASD related to its sensitivity in detecting various levels of language performance. Results indicated that Total Language raw scores were strongly associated with nonverbal mental age (regression coefficient = .85), which was measured by the Merrill-Palmer-Revised Scales of Development (Roid & Sampers, 2004). A floor effect was noted with only 30% of the sample, in which participants performed achieved a Total Language score of 50 (mean = 100; SD = 15). In the present study, a floor effect was only observed with one participant, Dylan, who was in the experimental group. The current study also found a strong relationship between expressive language skills and overall level of developmental functioning, which is consistent with the results by Volden et al.

Despite evidence that the PLS-4 is a suitable language assessment tool for children with ASD, it may not have been sensitive enough to detect changes in language for the purpose of the current study. This is consistent with research by Magiati and colleagues (2007), who documented improvements in age-equivalencies but not in standard scores on standardized tests measuring language abilities in preschool-aged children with ASD receiving home-based EIBI for two years. Instead, number of trials to mastery for EIBI tasks during baseline versus when RIRD was in effect may have been a better measure of progress
VOCAL STEREOTYPY

for expressive and receptive language skills over time. This finding supports the use of criterion-based instruments, such as the Assessment of Basic Language and Learning Skills-Revised (ABLLS-R; Partington, 2008) and Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008), in addition to norm-referenced tools when monitoring language gains in this population (Gould et al., 2011).

The Relationships between Language Skills, Developmental Level, and Vocal Stereotypy

It was hypothesized that increased standard scores on receptive and expressive language subscales on the PLS-4 and would be negatively correlated with percentage of vocal stereotypy. This hypothesis was not supported. Results indicated that there was no relationship between vocal stereotypy level for the last 5 sessions and developmental functioning, auditory comprehension, or expressive communication skills for participants in the experimental and clinical control groups. It was further hypothesized that developmental functioning level would be positively associated with receptive and expressive language skills. This hypothesis was partially supported. Specifically, a strong positive relationship was found between expressive communication skills at post-testing and developmental level. No relationship was found between auditory comprehension skills and overall developmental level. However, although not supported by these findings, given this study’s small sample size, it is still possible that Steve may have responded more favorably than Dylan to RIRD because he had a higher developmental level of functioning at the start of the study (MSEL Early Learning Composites = 60 versus 49, respectively).

Some research has evaluated the predictors for success in EIBI programs. Remington and colleagues (2007) found that children who responded positively to behavioral
intervention compared to nonresponders had a higher IQ, mental age, and language and social skills scores on the Vineland Adaptive Behavior Scales-Second Edition (Sparrow, Balla, & Cicchetti, 1984). In addition, Ray-Subramanian and Weismer (in press) found an inverse relationship between motor stereotypy and language skills (i.e., receptive and expressive) in young children with ASD. Sherer and Schreibman (2005) examined archival behavioral assessment data for 28 children who received pivotal response training (PRT), a child-led ABA intervention aimed to increase communication, social, and play skills in children with ASD. Through this analysis, the authors developed behavioral profiles associated with favorable treatment outcomes after identifying six participants as “exceptional” or “poor” responders to PRT. Exceptional responders to PRT had higher levels of stereotypic language and toy engagement and less social avoidance (i.e., they tolerated having another person within close proximity). Moreover, this behavioral profile accurately predicted which children would have better therapy outcomes for a different cohort of children receiving PRT. This study suggests that vocal stereotypy may not have interfered with acquisition during EIBI in the present study. Additionally, the results of the Sherer and Schreibman (2005) study are consistent with Lovaas and colleagues (1973), in which children with echolalia made more overall gains during the course EIBI.

Howlin, Magiati, and Charman (2009) conducted a recent review of outcome predictor variables for EIBI for children with ASD across multiple key studies. Their findings indicated that initial cognitive functioning level only sometimes predicted positive outcomes (e.g., Eldevik et al., 2006; Lovaas, 1987; Magiati et al., 2007; McEachin et al., 1993), whereas other author found a diagnosis of Pervasive Developmental Disorder Not Otherwise Specified rather than Autistic Disorder was predictive of favorable response to
EIBI (e.g., Smith et al., 2000). Moreover, some studies did not analyze predictor variables (e.g., Cohen et al., 2006; Howard et al., 2005). Operational definitions for “positive outcomes” were unavailable or unclear across studies. Given the mixed results of the Howlin et al. (2009) review, participant characteristics associated with better treatment outcome continue to be poorly understood and more research is needed in this area, as well as how the field should define favorable intervention outcomes (Matson, Tureck, Turygin, Beighley, & Rieske, 2012).

**Future Research Directions**

There are many reasons that necessitate continued research in the area of vocal stereotypy. In particular, vocal stereotypic behaviors are difficult to treat. The potentially cumbersome implementation of RIRD is an important factor for practitioners to consider when selecting a treatment to reduce vocal stereotypy. For instance, Miguel and colleagues (2009) reported that RIRD was implemented more than 100 times across an entire day to reduce vocal stereotypy in their participant’s natural classroom environment. Although Liu-Gitz and Banda (2010) reported positive findings when applying RIRD within a special education classroom setting, the RIRD procedure may not be feasible for caregivers and school personnel to implement across the day. Ahrens and colleagues (2011) demonstrated that RIRD had limited effectiveness in suppressing vocal stereotypy when the procedure is implemented intermittently (e.g., 10% or 25% of opportunities). As such, if providers, educators, and caregivers are not able to implement RIRD with high procedural fidelity, this intervention may not maintain to the natural setting. Negative side effects have also been noted during RIRD, such as emotional responding and physical aggression (Cassella et al., 2011). Therefore, additional research is warranted for developing effective treatment
packages for vocal stereotypy, while minimizing the likelihood of emergence of other problem behavior.

No studies have investigated the effects of RIRD within a treatment package, such as in combination with a DRO component. As RIRD is a punishment-based procedure, reinforcement-based behavior-reduction contingencies should be utilized whenever possible to increase a child’s access to reinforcers. When RIRD was implemented during EIBI with Dylan and Steve, the delivery of reinforcement was delayed because learning trials were suspended until participants met the termination criteria of three correct vocal or motor imitations without engaging in vocal stereotypy. These criteria were selected to be consistent with procedures outlined by Ahearn and colleagues (2007). However, some authors have utilized less stringent protocols for discontinuing RIRD trials when implemented contingent on vocal stereotypy. Cassella et al. (2011) terminated motor RIRD trials following three consecutive responses without vocal stereotypy, even if these responses were physically prompted (i.e., the participant made an error or did not respond within 5 seconds when given an instruction). Ahrens et al. (2011) demonstrated that compliance to vocal tasks was not necessary to achieve decelerative effects for vocal stereotypy; instead, therapists continued the RIRD sequence until the participant was exposed to three tasks without engaging in vocal stereotypy and praise was delivered contingent on independent or prompted correct responses. These data suggest that compliance to RIRD tasks is not a necessary condition for this intervention to be effective.

Given the mixed findings of the current study, direct replications are warranted. It is difficult to draw conclusions regarding the impact of RIRD on language skills, as measured by standardized language assessments, because RIRD was not successful in reducing vocal
stereotypy levels for one of two participants. In addition to utilizing repeated standardized language measures over time, data should be collected on the number of trials to mastery across programs to determine if the participant acquired target skills more quickly when RIRD was in place. As strategies for identifying effective RIRD tasks have yet to be developed, investigating the effects of RIRD tasks delivered during EIBI programs that require a vocal response (e.g., tacting, intraverbal) versus programs requiring pointing response (e.g., receptive identification, matching to sample) would be worthwhile. Types of tasks during RIRD also could be manipulated (e.g., vocal versus gross motor imitation) to identify the best combination for use during EIBI. During implementation of RIRD during an EIBI program, it is unknown if the topography of RIRD task (i.e., vocal versus nonvocal) should be different from the EIBI task. Similar to findings by Ahrens and colleagues (2011), RIRD tasks requiring a nonvocal response may be more favorable overall because physical guidance could be implemented when the child does not comply; however, when motor RIRD trials were introduced for Dylan, no decelerative effect was observed.

Future research could also evaluate two or more effective treatments for vocal stereotypy within a combined reversal and multielement design to determine the best method for an individual. Several factors should be examined, including the duration of procedures, side effects (e.g., crying, negative vocalizations, physical aggression, etc.), procedural fidelity, and generalizability to the natural environment. Exploring the use of contingency correlated-stimuli during RIRD may be beneficial for establishing stimulus control and teaching children to discriminate when procedures are in effect. For instance, Schumacher and Rapp (2011) used a large piece of red construction paper as a discriminative stimulus to signal when the RIRD contingency was in place. Evaluating other types of portable
discriminative stimuli also would be useful, especially in the contexts of caregiver and teacher training.

Overall, there is a paucity of behavior analytic research on empirically-supported treatments (ESTs) for children with ASD and intellectual disabilities with regard to behavior-reduction procedures. Currently, RIRD and DRO have the most empirical support for treating vocal stereotypy (Lanovaz & Sladeczek, 2012). Chambless and Hollon (1998) published several criteria for establishing efficacy for ESTs. These criteria include: (1) clearly defined samples determined by rigorous diagnostic assessment, (2) sound research methodologies that can be replicable, (3) use of treatment manuals, (4) valid and reliable assessment methods to measure outcome, (5) inclusion of follow-up data, and (6) data analysis of important factors such as attrition rate and therapist effects. Based on empirical level of support, ESTs are divided into three categories: well established, probably efficacious and possibly efficacious (Chambless et al., 1998). For the highest tier, well-established, interventions must have been demonstrated to be superior to another treatment for a well-defined sample (defined by diagnosis and demographic factors), across multiple studies and at least two independent research laboratories, and follow a treatment manual (Chambless et al., 1998). Generally, EST research incorporates the use of randomized-controlled trials (RCTs; Joyce, Wolfaardt, Sribney, & Awylin, 2006) but other between-group (e.g., quasi-experimental) and single-subject experimental designs (e.g., multiple-baseline across participants, reversals) are acceptable for the well-established category. Single-subject design studies can meet criteria with the inclusion of at least 9 participants, provided that the target intervention is compared to at least one other treatment and experimental control is demonstrated (Carr, Severton, & Lepper, 2009). To qualify as
probably efficacious, criteria regarding rigorous experimental design, clearly defined samples, comparison to another treatment or a waitlist control, and use of treatment manuals are still required but only three participants are needed for single-subject designs (Chambless et al., 1998; Chambless & Hollon, 1998). Lastly, interventions that have been effective for at least three participants using a replicable experimental design and no documentation of contradictory findings are classified as possibly efficacious.

Using Chambless and Hollon’s criteria, none of the ABA treatment studies for vocal stereotypy meet the requirements for well-established or probably efficacious. RIRD (Ahearn et al., 2007; Ahrens et al., 2011) and tact correction (Karmali et al., 2005) are the only treatments for vocal stereotypy that would meet criteria for possibly efficacious with the inclusion of three to five participants and use of acceptable single-subject experimental designs. RIRD has been demonstrated as effective across independent laboratory sites with this sample size (i.e., Ahearn et al., 2007; Ahrens et al., 2011). However, results from the current study indicate that RIRD was unsuccessful in decreasing Dylan’s vocal stereotypy and variables regarding ineffectiveness are unknown. In addition, for RIRD to meet criteria as a probably efficacious treatment in the future, participant samples must be clearly delineated. Specifically, diagnoses of ASD should be confirmed by an independent evaluator with gold standard assessment tools such as the ADOS and Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 1994). Treatment manuals regarding implementation of RIRD need to be developed and incorporated into studies for RIRD to qualify as probably efficacious. Smith, Scahill et al. (2007) recommended that manuals be written for treatment packages after initial efficacy studies for each specific procedure are conducted within single-subject designs. Following development of the manual, social acceptability and
clinician fidelity of treatment procedures should be evaluated prior to applying it within an RCT or other between-groups design. Thus, prior developing an RIRD manual, considerably more research is needed regarding the necessary and sufficient conditions for establishing efficacy within an RIRD package.

To better understand the trajectory of ASD symptoms in the absence of a specific treatment, group studies with randomized or quasi-experimental designs are needed with the inclusion of comparison groups, such as the control groups used in the present study with a larger number of participants (Shadish et al., 2002). For vocal stereotypy intervention research within a group design, children could be matched on a variety of variables, including chronological age, sex, expressive language abilities, adaptive skills, cognitive functioning level, or symptom severity based on nomothetic diagnostic tools. However, using RCTs to evaluate ABA interventions is a controversial issue (see Keenan & Dillenburger, 2011). Smith (2012) advocated the usage of between-group designs in behavior-analytic research for individual procedures following repeated demonstration of effective change on behavior; however, he urged that RCTs be the final step in verifying the efficacy of an intervention, consistent with previous recommendations by Johnston (1988). Even well-studied ABA interventions such as functional communication training (FCT; Carr & Durand, 1985) are not prepared for evaluation during an RCT because no general consensus exists regarding the standard FCT protocol, and there is limited research on procedures for thinning reinforcement schedules and programming generalization (Fisher, 2012). Hence, given that vocal stereotypy continues to be a challenging behavior to reduce and research on RIRD is still emerging, behavior analysts have a long journey before investigating vocal stereotypy interventions within large-scale studies with group designs.
Limitations

With respect to response measurement, vocal stereotypy data were collected using a 10-second PIR system. Limitations have been noted with PIR, such as underestimating the rate of high-frequency behaviors, overestimating the total duration of behavior, and requiring observers to monitor the behavior continuously throughout a session (Cooper et al., 2007). A common alternative for collecting data on vocal stereotypy includes momentary time sampling (MTS). MTS is an event recording method in which the response is recorded only if it occurs exactly at a predetermined moment (Harrop & Daniels, 1986). This procedure has been shown to have superior accuracy in estimating the actual duration of vocal stereotypy (Gardenier et al., 2004; Meany-Daboul, Roscoe, Bourret, & Ahearn, 2007) and does not require data collectors to monitor the target behavior outside of the intervals specified for time sampling. In recent research on RIRD, some authors (e.g., Ahearn et al., 2007; Cassella et al., 2011; Colon et al., 2012) have employed MTS in favor of other forms of interval recording, including PIR and whole-interval recording. However, despite the advantages of MTS, PIR was selected as the measurement method in the current study because the experimenter did not have access to a computerized data collection system with capabilities for MTS.

The fact that children in the experimental and clinical control groups experienced only 6 hours of EIBI per week deserves attention because this level of service is considerably less intensive than current recommendations. To achieve optimal results, at least 30 to 40 hours weekly of behavioral intervention is generally recommended for young children with ASD for at least 2 years (Eikeseth, 2009). In a recent review on EIBI practices for young children with ASD, Matson and colleagues (2012) reported that the majority of EIBI
programs offer 20 to 40 hours of therapy per week. Consistent with these guidelines, Granpeesheh, Dixon, Tarbox, Kaplan, and Wilke (2009) found that children with ASD between 2 to 7 years of age were more likely to benefit from increased hours of EIBI compared to older children, as measured by the number of mastered goals across time. In addition, Eldevik, Eikeseth, Jahr, and Smith (2006) found that while there were statistically significant differences in favor of EIBI between two groups of children who received EIBI or eclectic therapy for 12 hours per week, these discrepancies were small and may not have been clinically meaningful. As such, it is possible that results of the present study may have been different if children within the experimental and clinical control groups received behavioral intervention at the intensity commensurate with current practice recommendations. Nevertheless, when taking into account family and clinic resources as well as the maximum amount of weekly hours reimbursable by medical insurance, it would have been difficult to provide EIBI services at this level of care for the purposes of the current investigation.

A methodological weakness was that the MSEL, ADOS, and PLS-4 (pre-testing) assessments were conducted by the experimenter, as well as the PLS-4 post-testing for three participants. Additionally, the clinician who administered the PLS-4 post-testing to the other three participants was not blind to study hypotheses. This suggests that testing could have been influenced by experimenter bias. The experimenter did not use alternative examiners due to lack of available personnel as there were no graduate students or therapy staff trained in psychodiagnostic test administration at one of the recruitment facilities. A stronger design would have been to include an independent evaluator, blind of study procedures and hypotheses, to conduct all assessments in order to reduce the likelihood of bias.
Another limitation was that randomization across groups not possible, given that all participants were already enrolled in some type of therapy. Due to the lack of randomization across groups, it is not possible to determine if participants in a particular group would have responded more favorably to a different intervention. The possibility that unidentified differences existed across groups cannot be ruled out, which could account for the changes observed in pre and post PLS-4 standard scores, especially with regard to the improvements in performance for participants within the traditional treatment control group. For instance, families who chose to enroll their child into EIBI instead of alternative, traditional therapies (e.g., occupational therapy) may have differed on demographic variables, such as parental educational level or other variables unknown to the experimenter, which may have influenced participant therapeutic outcomes over time. While formal data on family demographic variables were not collected during this study, previous research has not found a relationship between demographic variables and therapy outcomes for children with ASD. In the Remington et al. (2007) study, mothers who enrolled their children in EIBI were more likely to have a college education, yet differences to mothers of children in the comparison group were not statistically significant. Similarly, Miller, Schreck, Mulick, and Butter (2012) found that parent level of education, annual income, and age were not associated with choosing an empirically-supported treatment approach. Rather, parents’ selection of treatments for their children with ASD was most influenced by recommendations from school personnel and other parents, even though educators, speech-language pathologists, occupational therapists, and physical therapists were least likely to suggest interventions supported by science. However, it remains unclear if family demographic variables impacted the results of the current study.
References


Erlbaum Associates.


VOCAL STEREOTYPY


Tucker, D. J., & Berry, G. W. (1980). Teaching severely multihandicapped students to put
on their own hearing aids. *Journal of Applied Behavior Analysis, 13*(1), 65-75.


Appendix A: Eastern Michigan University Human Subjects Approval Letter

EASTERN MICHIGAN UNIVERSITY  
Education First

August 17, 2010

To: Tammy Perry  
Psychology

Re: UHSRC # 100701  
Approval Date: August 10, 2010

Title: The Effects of Language Skills and Discrete Trial Training on Response Interruption and Redirection as Treatment for Vocal Stereotypy

The Eastern Michigan University Human Subjects Review Committee (UHSRC) has completed their review of your project. I am pleased to advise you that your research has been approved in accordance with federal regulations.

Renewals: Full Board review protocols need to be renewed annually. If the project is continuing, please submit the Human Subjects Continuation Form prior to the approval expiration. If the project is completed, please submit the Human Subjects Study Completion Form (both forms are found on the UHSRC website).

Revisions: Full Board review protocols do require revisions. If changes are made to a protocol, please submit a Human Subjects Minor Modification Form or new Human Subjects Approval Request Form (if major changes) for review (see UHSRC website for forms).

Note that all requests for modification and continuation require a full board review. Forms need to be submitted at least one month in advance of the approval expiration to allow time for review.

Problems: If issues should arise during the conduct of the research, such as unanticipated problems, adverse events, or any problem that may increase the risk to human subjects and change the category of review, notify the UHSRC office within 24 hours. Any complaints from participants regarding the risk and benefits of the project must be reported to the UHSRC.

Follow-up: If your Full Board review project is not completed and closed after three years, the UHSRC office will require a new Human Subjects Approval Request Form prior to approving a continuation beyond three years.

Please use the UHSRC number listed above on any forms submitted that relate to this project, or on any correspondence with the UHSRC office.

Good luck in your research. If we can be of further assistance, please contact us at 734-487-0042 or via e-mail at human.subjects@emich.edu. Thank you for your cooperation.

Sincerely,

Deb de Laski-Smith, Ph.D.  
Interim Dean  
Graduate School  
Administrative Co-Chair  
University Human Subjects Review Committee

University Human Subjects Review Committee · Eastern Michigan University · 200 Boone Hall  
Ypsilanti, Michigan 48197  
Phone 734.487.0042  Fax 734.487.0060  
E-mail human.subjects@emich.edu  
www.ord.emich.edu (see Federal Compliance)

The EMU UHSRC complies with the Title 45 Code of Federal Regulations part 46 (45 CFR 46) under FWA00000050.
Appendix B: University of Nebraska Medical Center Human Subjects Approval Letter

December 6, 2011

Tamara Leigh Perry, MS  
Center for Autism Spectrum Disorders  
UNMC - 5450

IRB # 620-11 FB

TITLE OF PROPOSAL: The Effects of Response Interruption and Redirection on Language Skills in Children with Vocal Stereotypy

SECONDARY INVESTIGATOR: Michael E. Kelley, PhD.

DATE OF FULL BOARD REVIEW: 11/22/2011

DATE OF FINAL APPROVAL: 12/06/2011  VALID UNTIL 11/22/2012

SUBPART B, C, AND/OR D CATEGORY OF REVIEW: D - 404

The Institutional Review Board (IRB) for the Protection of Human Subjects has completed its review of the above-titled protocol, and informed consent documents, including any revised material submitted in response to the IRB's review. The IRB has expressed it as their opinion that you are in compliance with HHS Regulations (45 CFR 46), applicable FDA Regulations (21 CFR 50, 56) and the institution's HRPP Policies and you have provided adequate safeguards for protecting the rights and welfare of the subjects to be involved in this study. This letter constitutes official notification of the final unconditional approval and release of your project by the IRB, and you are authorized to implement this study as of the above date of final approval.

Please be advised that only the IRB approved and stamped consent/assent form can be used to make copies to enroll subjects. Also, at the time of consent all subjects/legally authorized representatives (LARs) must be given a copy of The Rights of Research Subjects and "What Do I Need to Know" forms. The IRB wishes to remind you that the principal investigator is responsible for ensuring that ethnically and legally effective informed consent has been obtained from all research subjects.

Finally, under the provisions of this institution's Federal Wide Assurance (FWA00002939), the principal investigator is directly responsible for submitting to the IRB any proposed change in the research or the consent/assent document(s). In addition, any adverse events and unanticipated problems involving risk to the subject or others must be promptly reported to the IRB.

In accordance with HRPP Policy, this project is subject to periodic review and surveillance by the IRB and, as part of their surveillance, the IRB may request periodic progress reports. For projects which continue beyond one year, it is the responsibility of the principal investigator to initiate a request to the IRB for continuing review and update of the research project.

Sincerely,

Bruce G. Gordon, MD  
Chairman, Joint Pediatric IRB

BGG: sah
Appendix C: Informed Consent Form

“The Effects of Language Skills and Discrete Trial Training on Response Interruption and Redirection as Treatment for Vocal Stereotypy”

To be conducted by Tamara Perry, MS, BCBA (Doctoral Candidate) and James Todd (Professor of Psychology)

Eastern Michigan University

1. **Purpose of Research Study**: The purpose is to measure the effects of a language program on reducing inappropriate language (“vocal stereotypy”) and improving language skills in children with autism spectrum disorders. If your child is currently receiving or will be receiving applied behavior analysis (ABA) therapy, he or she will be randomly assigned to our ABA group or our ABA-plus-language program group. During the ABA-plus-language group, participants will receive an additional language program to replace inappropriate language, which would occur during your child’s regular treatment.

If you do not want your child to participate in ABA at this time, you may have the option of your child participating in our “treatment as usual group.” In this case, your child would continue with his or her usual treatment plan (e.g., music therapy, occupational therapy, etc.) and you would be asked to release the results of your child’s testing and complete additional language testing in approximately 3 months from signing this form. The purpose of the “treatment as usual” group is to provide a comparison group to evaluate the effects of the language program.

2. **Participation Withdrawal or Refusal to Participate**: Participation in this study is purely voluntary. Refusal to participate will not result in any penalty or a loss of benefits. You may decide to withdraw your child from the study at any time without consequences to your child’s current or future treatment.

3. **Participation Requirements**: Your child must be between the ages of 3 and 5 years old at the start of the study and have a diagnosis of an autism spectrum disorder in order to participate. Basically, we will conduct language and diagnostic testing of your child that are already a part of your child’s entry into the treatment program, using tests that are commonly used with children in clinics and schools. Before your child’s participation may begin, you will be asked to sign this consent form to verify that you understand the study procedures and are willing to allow your child to work with the principal investigator. We will not ask you to sign this document if it appears you do not understand it or the relevant aspects of the study.

4. **Description of Study Procedures**: If your child is assigned to in the ABA-plus-language program, the study will include 4 parts:
   a. An *assessment phase* in which your child will receive developmental, diagnostic, and language testing to find out his or her current functioning. This testing will be completed as part of your child’s routine treatment at the agency that your child currently attends or plans to attend.
   b. A “*functional analysis,***” in which the cause of your child’s identified language problems is investigated during brief sessions of about 5 minutes by watching your child’s reactions to toys or the actions of others. The experimenter will systematically
change how she acts to determine when your child is most likely to use inappropriate language (e.g., repeating movie lines, making animal noises).

c. **A treatment phase**, which has several parts: (1) Baseline: Your child’s inappropriate language will be measured during academic tasks for 5 minutes. Several 5-minute sessions will be conducted to determine a pattern. (2) Treatment: Whenever the child uses inappropriate language, the therapist will stop the current task and tell the child to complete 3 language tasks correctly. For example, the child may be asked to label pictures, answer simple questions (e.g., “What is your name?”), or say words that he or she already knows. This procedure will be repeated until the child stops using inappropriate language. Once the child is no longer engaging in inappropriate language at that moment, the current ABA program will resume. All sessions will be conducted in a small room at your child’s treatment center with the experimenter present, and last approximately 5 minutes each, for no longer than 60 minutes per day. Treatment will last for approximately 2 months.

d. **A language re-testing phase**, in which the language test that was given to your child at the beginning of the study will be repeated approximately 3 months later. The purpose of this assessment will be to see if any improvements in your child’s language skills have occurred following the language program or any other program that your child has participated in during the course of the study. This language test should take approximately 30 to 60 minutes to complete and will be free of charge.

If your child is assigned to the ABA group, he or she will participate in all of the above study procedures except for Part C, the treatment phase.

If you do not want your child to participate in the treatment part of this study, but are willing to give permission for in the “treatment as usual” group, you will be asked to release your child’s test results and complete only Part D (above) of the study procedures listed.

5. **Possible Benefits of Participating:** If your child participates in the ABA-plus-language program, he or she may benefit from the considerable experience shared by the investigator, Ms. Perry, and her mentor, Dr. James Todd, who have worked with children with autism spectrum disorders for many years. If the program that we develop for your child is successful, it is possible that you will see reductions in your child’s inappropriate vocalizations as well as an increase in new language skills. Reducing inappropriate language may also allow for increases in other social skills, help the child become more independent, and improve relationships with peers. Alternatively, it is possible that there will be no change. Regardless of whether your child participates in the ABA-plus-language program, you will learn valuable information about your child’s language progress through the free language testing that will occur at the beginning of the study and three months later.

6. **Possible Risks of Participating:** The risks associated with this study are unlikely to exceed those that are ordinarily expected from your child’s current therapy. Precautions will be taken to minimize risks to your child and all sessions will be conducted in a safe environment that is free of any hazardous materials. The researcher will follow all safety and behavior management policies in accordance with the treatment center’s overall policies that are in place outside of the research context. Emergency situations (e.g. seizures, major injuries due to behavior problems, or other accidents) will be handled by taking appropriate steps to maintain the child’s safety and the safety of others until the guardian is able to arrive. Parents will be notified immediately of any problems. All appropriate health and safety regulations will be followed. If unexpected behavior issues arise during the study, or there are other health or safety problems that might prevent your child’s continued participation, the
situation will be discussed with you and others as appropriate in a timely manner before participation continues. In the unlikely event of an injury resulting from the research study, no reimbursement, compensation, or free medical care is offered by the researchers conducting the study.

7. **Usage and Storage of Research Results**: The results of the study will become Ms. Perry’s Doctoral Dissertation, and may be published in journals or other academic outlets, discussed as cases in college-level courses, or presented at conferences. Your child’s participation will be kept confidential, and any presentation or publication of the findings will not identify you or your child. When presenting this data, your child will be de-identified using an alternate name or number. All data will be physically stored in a locked file cabinet by the principal investigator, with all identifying information removed. All testing protocols will be kept at the child’s treatment center in his or her regular file and will not leave the building. Data stored on computers will also be de-identified using an alternate name for your child.

8. **Alternative Treatments Outside of the Study**: Researchers have identified several procedures, including the one under investigation, to minimize the use of inappropriate language. These treatments include ignoring the behavior, rewarding opposite behavior, and repeating appropriate language until the child stops saying inappropriate words or sounds. Caregivers may choose to explore alternative options outside the context of this study. Other behavior reduction or medical treatments will not be offered as part of the current study. Should any new findings regarding the procedures under investigation come up during the course of this study, Ms. Perry will discuss them with you as appropriate.

9. **Study Contact Information**: Any questions or comments about the study may be directed to Tamara Perry (734-635-8655/tpawich@emich.edu) or Professor James Todd (734-487-0376/jtodd@emich.edu). Dr. Todd may also be written at: Eastern Michigan University, Department of Psychology, 537 Mark Jefferson Hall, Ypsilanti, Michigan, 48197.

This research protocol and informed consent document has been reviewed and approved by the Eastern Michigan University Human Subjects Review Committee for use from 8/10/11 to 8/10/12. If you have questions about the approval process, please contact Dr. Deb de Laski-Smith (734.487.0042, Interim Dean of the Graduate School and Administrative Co-Chair of UHSCR, human.subjects@emich.edu).

If you have read and understood the above and will give your child permission to participate, please provide your name, date, and signature below. By doing so, you are giving informed and voluntary consent.

Please check the following:

- I agree for my child to participate in the study.
- I agree for my child to participate in “treatment as usual” group ONLY and allow release of his/her test results to the principal investigator.

___________________________
Your Child’s Name

___________________________
Parent’s Name (Please Print)

___________________________
Parent’s Signature/Date

___________________________
Signature of Principal Investigator/Date
Appendix D: Steve’s Data Sheet

Condition: _________ Participant #: _____ Observers: _____________ Date: ___________

Write a “V” in each interval box that the participant engages in vocal stereotypy at the end of the interval when the CD says, “Record.” Write a “—” in each interval box that vocal stereotypy does not occur at the end of the interval. All intervals are 10 seconds in duration and sessions last for 5 minutes.

<table>
<thead>
<tr>
<th>Session #:</th>
<th>% Vocal Stereotypy:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session #:</th>
<th>% Vocal Stereotypy:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>