

Evaluating the Use of a Multiple Schedule as a Method for Identifying Treatment Effects

**A Thesis Presented**

**By**

**Lynn Andrejczyk**

The Department of Counseling and Applied Educational Psychology

In partial fulfillment of the requirements

for a degree of

**Master of Science**

**in the field of**

Applied Behavior Analysis

Northeastern University

Boston, MA

August 2010

UMI Number: 1482876

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 1482876

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

NORTHEASTERN UNIVERSITY

Bouv  College of Health Science Graduate School

Thesis Title: Evaluating the Use of a Multiple Schedule as a Method for Identifying  
Treatment Effects

Author: Lynn Andrejczyk

Department: Counseling and Applied Educational Psychology

Approved for Thesis Requirements of Master of Science Degree

---

William Ahearn

---

Eileen Roscoe

---

Susan Langer

Evaluating the Use of a Multiple Schedule as a Method for Identifying Treatment Effects

By

Lynn Andrejczyk

B.A., Saint Anselm College

Submitted In partial fulfillment of the requirements for the degree of Master of Science in Applied Behavior Analysis in Bouvé College of Health Sciences Graduate School of Northeastern University, August 2010

## Acknowledgements

I would like to thank The New England Center for Children for allowing me to conduct my research and the participants for their cooperation. I would like to thank Andrea Carey for running sessions and scoring data. I would also like to thank my thesis committee, Bill Ahearn, Eileen Roscoe, and Susan Langer for their assistance in helping me through this process.

## Table of Contents

### I. Evaluating the Use of a Multiple Schedule as a Method for Identifying Treatment

#### Effects

A. Abstract .....	2
B. Introduction .....	3
1. Functional Analysis .....	3
2. Preference Assessment .....	6
3. Stereotypy and Treatment .....	9
4. Multiple Schedule .....	15
5. Purpose .....	17
C. Method .....	18
1. Participants, Settings & Materials .....	18
2. Response Measurement and IOA .....	19
3. Procedure .....	20
4. Experimental Design .....	22
D. Results.....	23
E. Discussion .....	26
F. References .....	29
G. Tables .....	32
H. Figure Captions .....	33
I. Figures .....	34

### Abstract

Many traditional methods for evaluating behavior (e.g., multiple baseline, alternating treatments, withdrawal, etc.), while effective, can often prove to be time and labor intensive. The purpose of this study was to evaluate the use of a brief methodology for identifying treatment effects. Participants were 2 children diagnosed with an autism spectrum disorder who exhibited high levels of vocal stereotypy. For both participants, functional analyses showed that the behavior was likely maintained by automatic contingencies of reinforcement. Following the functional assessment, vocal stereotypy was exposed to baseline, motor response interruption and redirection (RIRDm), and vocal response interruption and redirection (RIRDv) conditions using a multiple schedule design. RIRDm and RIRDv procedures were similar to the procedures reported by Ahearn, Clark, MacDonald, and Chung (2007). Following the multiple schedule, vocal stereotypy was exposed to the same conditions using a traditional withdrawal design (ABABACAC). A proportional analysis of the data was completed in order to determine whether a difference was obtained between the baseline and treatment conditions. Results showed that use of the multiple schedule produced a consistent treatment effect for one of the participants. Both RIRDm and RIRDv produced lower levels of behavior in treatment relative to baseline. A similar treatment effect was obtained when exposing that participant's vocal stereotypy to both treatments within the traditional withdrawal design.

## Evaluating the Use of a Multiple Schedule as a Method for Identifying Treatment Effects

### **Introduction**

In the applied setting, it is important that clinicians have the ability to move from behavioral assessment to treatment as effectively and efficiently as possible. Several studies have evaluated the use of brief assessments while providing intervention, including brief functional analyses and preference assessments. Many of these studies have effectively adapted these technologies to be conducted in a more efficient manner without significantly compromising the validity or integrity of results, allowing clinicians to move quickly from behavioral assessment to treatment evaluation.

While there are many methods with which one can effectively and confidently evaluate treatments, there have yet to be many attempts to develop and evaluate the use of a brief treatment evaluation method in the applied setting. While effective, many of the traditional methods (e.g., multiple baseline, alternating treatments, withdrawal, etc.), can become a time consuming and laborious process. Such analyses also require significant resources and may delay the implementation of effective behavioral treatment.

### **Functional Analysis**

Iwata, Dorsey, Slifer, Bauman, and Richman (1994/1982) standardized the use of the multi-element design to expose individuals to conditions for the purpose of identifying the functional cause of problem behavior when conducting an analog functional analysis. The target behavior they chose to assess was self-injury, and conditions consisted of (1) social disapproval (attention) in which attention was provided contingent upon the occurrence of the target behavior, (2) academic demand (escape) in



which a demand was presented and then removed for a brief period contingent upon occurrence of the target behavior, (3) unstructured play in which the participant had free access to a variety of toys and the experimenter would deliver brief social praise contingent upon appropriate behavior, and (4) alone in which the participant was left alone in the therapy room, without access to toys or other materials, and all behaviors were ignored. All sessions were 15 min in length, unless predetermined criteria to discontinue the session was met. The analysis continued until stable rates of behavior were obtained, unstable levels continued in all conditions, or 12 days of sessions were completed. The length of subject participation averaged 8 days or 30 sessions. If the highest levels of behavior were observed during the academic demand condition, it implied that social negative reinforcement (escape) maintained problem behavior. If highest levels of the behavior were observed during social disapproval, it implied that social positive reinforcement (attention) maintained problem behavior. If high levels of the behavior were observed during the alone condition, it was implied that non-social or automatic reinforcement maintained problem behavior. Of the nine participants in their initial study, a distinct maintaining contingency was identified for six.

In 1994, Iwata, Pace, Dorsey, Zarcone, Vollmer, Smith et al., published an extension of this study, in which 152 participants' self-injury was assessed using the above procedure. Of all the cases included, only 5% produced undifferentiated or inconclusive results.

In 1992, Derby et al. adapted functional analysis procedures described by Iwata, et al. (1994/1982) to fit a 90-min time limited outpatient clinic. In this study, 79 cases were evaluated in a 3-year period. The assessment method they developed utilized the

conditions of Iwata et al. (1994/1982). The assessment consisted of a single session of each condition, presented in a multi-element design. Then, a replication phase, in which conditions resulting in the highest and lowest levels of behavior were repeated, was conducted. They sought to report on what percentage of clients the brief functional assessment identified a specific maintaining variable, and whether or not, through manipulating that maintaining variable, they could show a decrease in aberrant behavior. They also sought to evaluate the extent to which the assessment procedures were replicable across clients, staff, and response topographies.

Prior to each assessment, a morning staff meeting was held in which members of the clinical team would review the clients' history and develop a hypothesis regarding the maintaining conditions. During the 90-min clinic evaluation, each client would typically undergo an initial assessment, in which behavior was exposed to one, ten-minute session per condition. This was followed by a replication phase where a contingency reversal was implemented, in which contingencies maintaining aberrant behavior were provided for an appropriate response (i.e., manding). Participants were clients who had been evaluated by the Self-Injurious and Aggressive Behavior Service, Department of Pediatrics, The University of Iowa, from 1987 to 1990. Participants included 46 males and 33 females aged 1-year to 32 years, who were diagnosed with varying levels of mental retardation, or an unspecified disability.

Of the 63% of participants who displayed aberrant behavior during the 90-min evaluation clinic, a distinct maintaining condition was identified in 74% of cases. However, because only 63% of participants exhibited aberrant behavior during the brief

evaluation, this procedure may be most appropriate for problem behavior that occurs at a stable and high level in order to obtain reliable results.

### **Preference Assessments**

Preference assessment represents another area in which researchers have sought to decrease the amount of time it takes to obtain valid and reliable results. Since the development of the single-stimulus assessment (Pace, Ivancic, Edwards, Iwata, & Page, 1985), several methodologies have been designed to evaluate and identify preference more effectively and efficiently. Following the development of the single-stimulus assessment, Fisher, Piazza, Bowman, Hagopian, Owens, and Slevin (1992) developed a paired-stimulus forced-choice assessment, which resulted in greater differentiation among stimuli as well as better prediction as to which stimuli would function as reinforcers when presented contingently. Windsor, Piché, and Locke (1994) compared paired-stimulus and multiple-stimulus presentations of stimuli and found that preferences and stimulus ranking were comparable across formats, with consistency being higher in the paired-stimulus assessment. However, the paired-stimulus assessment generally took two to three times as long to administer as the multiple-stimulus assessment. One of the primary limitations of the multiple-stimulus presentation, was that when all stimuli were being presented during each trial, participants would often engage with their highest preferred item to the exclusion of the rest, therefore it was difficult to rank moderate and low preferred stimuli. DeLeon and Iwata (1996) addressed this in a study in which they compared the paired-stimulus, multiple-stimulus with replacement (MS), and multiple-stimulus without replacement (MSWO) formats.

In developing the MSWO, DeLeon and Iwata (1996) aimed to address this limitation of the MS, in order to make the more time efficient method as clear and informative as the lengthier paired-stimulus assessment. During the MSWO procedure, each session began with all stimuli available, and randomly sequenced in a straight line. The participants were instructed to select one item. Once an item was selected, it was either removed from the area (leisure items), or not replaced after being consumed (edible items). Prior to the next trial all remaining items were rotated by moving the item on the left to the far right end, and shifting the items so that they were equally spaced on the table. This procedure continued until all items had been selected or the participant made no selection within 30 s of the stimuli presentation. The MS procedure was similar to the MSWO however, instead of removing or not replacing selected or consumed items during subsequent trials, those items were replaced. During paired-stimulus sessions, two items were presented during each trial, and the session continued until each item had been paired with every other item in a predetermined order, with stimuli randomly positioned (left or right side).

Each participant was exposed to five consecutive sessions of each procedure, for a total of 15 sessions. The order of exposures to the procedures varied across participants. Results showed that, as in previous studies, multiple-stimulus procedures consistently took half the time it took to conduct the paired-stimulus assessment. Also, it was found the MSWO and paired-stimulus presentations produced more distinct and consistent rankings than the MS assessment. The MSWO was found to produce results comparable to those obtained with the paired-stimulus assessment in considerably less time.

Carr, Nicolson, and Higbee (2000), further extended research on the MSWO assessments in two ways. First they attempted to make the MSWO more time efficient by reducing the number of stimulus presentation arrays from five to three. Secondly they evaluated the effectiveness of their method in naturalistic contexts for children diagnosed with autism. There were three participants, 2-7 years of age, in the study. All of the participants attended a university-based day program, and all sessions were conducted in the participants' daily therapy rooms. Eight leisure and edible stimuli were selected for each participant from parent and therapist suggestion. A brief MSWO assessment and reinforcer evaluation was conducted for each participant.

MSWO procedures were identical to those described by DeLeon and Iwata (1996) with the exception that only three stimulus-presentation trials were conducted, rather than five. Following the MSWO assessment, a contingency evaluation was conducted for three of the stimuli: a high, medium and low preference item. A low-frequency target behavior was selected for each participant from his or her ongoing acquisition curriculum.

Results showed that for two participants the low-preference stimulus failed to significantly increase responding over baseline levels, and for one participant the medium-preference stimulus produced moderate reinforcement effects (Carr et al., 2000). The high-preference stimulus produced responding that was higher than baseline and medium and low-preference stimulus conditions for all participants. These results support the use of the brief MSWO assessment. In addition, both the brief MSWO assessment and reinforcer evaluation were conducted in less than 1 hr for each participant.

Paramore and Higbee (2005) replicated the Carr et al. (2000) study to the end of assessing the procedures' generality to adolescents diagnosed with emotional-behavioral disorders, as much of the previous research had focused on use of preference assessments in individuals with developmental disabilities such as mental retardation and autism (DeLeon & Iwata, 1996; Fisher et al., 1992; Pace et al., 1985). Procedures were identical to those described by Carr et al. with the following modifications: only edible stimuli were used in the stimulus array, the stimulus array consisted of five stimuli, instead of eight, and participants could indicate their selections either verbally or nonverbally. A reinforcer assessment was then conducted for each participant, targeting on-task behavior which was defined as sitting appropriately at the desk, working on the assigned task, and speaking only to ask task related questions after raising a hand and being called on by the teacher. The reinforcer assessment evaluated the use of a low-, medium-, or high-preference stimuli as a reinforcer when presented contingently for on-task behavior.

It was found that, although the results for the reinforcer evaluation were initially undifferentiated, ultimately all three participants' on-task behavior increased to highest observed rates when using the respective high-preference stimuli. With the exception of one participant, lowest rates of on-task behavior occurred when the low-preference stimulus was presented. These data not only replicate the findings of Carr et al. (2000), but also show that MSWO preference assessment can be generalized to other populations.

### **Stereotypy**

Repetitive or stereotyped behaviors, while not unique to individuals diagnosed with autism (Lewis & Bodfish, 1998), are one of its defining characteristics (Lewis & Bodfish; Rincover, 1978). Stereotypic behaviors can often interfere with other

prosocial/appropriate behaviors, such as discrimination learning and play, and may be socially stigmatizing (Rincover, 1978).

Stereotypic behaviors are considered to be operant behavior maintained by sensory consequences automatically produced by the behavior (Ahearn et al., 2003; Vollmer, 1994). Lovaas et al. (1987) explain this further stating that due to their often elaborate and idiosyncratic nature, learning variables must have shaped the stereotypic behavior as discrimination would be necessary to produce specific sensory consequences. Lovaas and colleagues goes on to explain that stereotypic behaviors can be described as automatically maintained because attention and other socially-mediated consequences can be introduced or withdrawn without causing an increase or reduction in the behavior, persisting indefinitely in the absence of social consequences. That stereotypic behavior occurs independently of social consequences is one of many difficulties in treating the behavior (Ahearn et al., 2007; Lovaas et al., 1987; Rapp & Vollmer, 2005; Vollmer, 1994).

Other difficulties encountered when treating stereotypic behavior include treatments that are often impractical or difficult to implement, as well as difficulty in interrupting the behavior itself (Lovaas et al., 1987). Because the behavior is automatically maintained, and it's consequences controlled by the individual emitting the behavior, stereotypy is often resistant to external consequences and processes such as extinction, punishment, and differential reinforcement (Vollmer, 1994). Despite such difficulties, treatments have been developed which were successful in decreasing rates of behavior maintained by automatic reinforcement. Sensory extinction, response blocking with or without the use of redirection, differential reinforcement, and interruption have all

been shown to successfully decrease rates of automatically-maintained behavior with or without the use of redirection (Fellner, LaRoche, & Sulzer-Azaroff, 1984; Hagopian & Adelinis, 2001; Lerman & Iwata, 1996; Rincover, 1978).

### **Treatment of Stereotypy**

In 1978, Rincover treated the stereotypic behaviors of three children ages 7-14, who had been diagnosed with autism or mental retardation. It was determined that the behaviors (i.e., spinning plate and listening, finger flapping in front of eyes, twirling objects between fingers in front of eyes) were automatically maintained and therefore Rincover sought to either block or mask the sensory consequences maintaining the behavior.

Sessions lasted 20 min and were conducted in a 2.5 X 2.5 m classroom, which held a table, chairs, and object the participant preferred to engage in stereotypy with. During sensory extinction sessions, different methods were employed to block the three types of sensory feedback that the behaviors produced. To eliminate the auditory feedback of the plate spinning, a carpet was installed on the table in the classroom. The carpet was such that it would not restrict the plate from spinning just from producing noise. To eliminate proprioceptive sensory feedback from finger flapping and object manipulation, a small vibratory mechanism was taped to the back of the participants' hand. The vibrator did not physically restrict stereotypic behavior. In order to restrict visual sensory feedback, a blindfold was introduced, consisting of a handkerchief, once folded, snugly placed over the participants' eyes and tied behind his or her head.

Baseline sessions were alternated with Sensory Extinction sessions in a reversal, multiple-baseline across participants design. During baseline, no differential



consequences were presented contingent upon stereotypy, and the objects the participants preferred to engage in stereotypy with were made available. During sensory extinction, the addition of a stimulus designed to eliminate the sensory feedback of the participant's stereotypy was introduced (i.e., carpet, vibrator, or blindfold).

Results showed that plate-spinning significantly decreased with the application of the carpet on the desk, and the vibrator significantly decreased finger flapping, and also decreased object twirling more effectively than use of the blindfold in comparison to baseline. While these procedures reduced the occurrence of stereotypic behavior, the methods are potentially impractical in the applied setting.

In 1984, Fellner et al. conducted a study to evaluate the use of differential reinforcement and interruption in decreasing rates of stereotypy. Their participant was a 6-year-old girl diagnosed with severe mental retardation. Thirteen responses were targeted for treatment: hand flapping, hyperventilating, slapping hand to mouth, lip pulling, finger flipping, pressing fingers together, rubbing her body, rocking upper body, gazing, pressing fingers to objects, scratching, hair pulling and finger staring. However, only hand flapping, hyperventilating, slapping hand to mouth, and lip pulling data were used to evaluate functional relations, as the other nine did not occur consistently throughout the day.

During sessions, the participant and therapist were seated side by side at a table, where a variety of toys were placed. An ABAB design was used to evaluate the participant's ongoing treatment, which was a combination of differential reinforcement of other behavior plus differential reinforcement of incompatible behavior (DRO plus DRI). Researchers also evaluated DRO plus DRI combined with an interruption procedure to

determine whether or not using interruption with differential reinforcement would be more effective in decreasing problem behavior than differential reinforcement alone.

In the DRO plus DRI condition, milk and toys that produced preferred sensory stimulation paired with social praise were used to reinforce alternate responses observed in the absence of the targeted responses. The initial DRO interval was 5 s, and was successively increased by 5 s each time the participant did not engage in target behaviors for three consecutive intervals. Reinforcement was also delivered contingent upon the occurrence of incompatible behaviors such as appropriate toy play.

During DRO, DRI, and interruption, the trainer would interrupt the targeted behaviors in addition to continuing the reinforcement contingencies. Interruption for hyperventilation entailed a verbal, “no” accompanied by the therapist cupping her hand over the participant’s mouth for 1-2 s. For the other targeted behaviors the therapist said, “hands down” and guided the participant’s hands to her lap. Results showed that DRO and DRI combined with interruption was more effective at decreasing rates of stereotypy than DRO and DRI alone.

Lerman and Iwata (1996) successfully implemented response blocking to decrease rates of hand mouthing in a 32-year-old man diagnosed with profound mental retardation. During baseline the participant was seated in a chair and no differential consequences were presented contingent upon behavior. During response blocking sessions, a therapist was seated behind the participant and would block some, or all occurrences of hand mouthing by placing the palm of her hand about 2 cm in front of the participant’s mouth. Rates decreased rapidly to near zero levels during response blocking sessions. Results of this study also suggested that response blocking suppresses behavior

through punishment rather than through the termination of the relevant reinforcer (extinction).

Hagopian and Adelinis (2001) successfully implemented response blocking with redirection to decrease rates of pica in a 26-year-old man diagnosed with moderate mental retardation and bipolar disorder. A functional analysis showed that pica was partly maintained by automatic reinforcement. Pica was defined as the participant attempting or successfully placing baited items (pieces of paper) or other inedible objects past the plane of his lips. For all sessions, the room was baited with small pieces of paper on the floor. Sessions were 5 min during the blocking analysis and 10 min during the treatment analysis. Because the participant typically emitted aggressive behavior when pica was blocked, experimenters first ran a blocking analysis, and then the treatment analysis.

During the blocking analysis a reversal of an ignore condition and response blocking was implemented. During ignore, no differential consequences were provided for pica or aggression. During response blocking, the therapist attempted to block all occurrences of pica by placing a hand between the participant's hand and mouth. No differential consequences were provided for aggression. During the treatment analysis, response blocking with redirection to an alternative food item (identified using a paired-choice preference assessment) was evaluated using a reversal design. Response blocking with redirection effectively decreased rates of pica, in the absence of aggression, in comparison to baseline, where response blocking alone did not decrease rates of pica, and caused an increase in aggression.

In 2007, Ahearn, Clark, MacDonald, and Chung used response interruption and redirection (RIRD) to decrease rates of vocal stereotypy. Vocal stereotypy was defined as any instance of nocontextual or nonfunctional speech and included singing, babbling, repetitive grunts, squeals, and phrases unrelated to the present situation. The participants were 2 boys and 2 girls who had been diagnosed with an autism spectrum disorder, and had been referred by their clinical and educational teams as exhibiting vocal stereotypy that interfered with their participation in instructional activities.

Effects of RIRD were evaluated in an ABAB design. Baseline sessions were 5 minutes in length and no differential consequences were provided for vocal stereotypy. If the student appropriately manded an available item, the therapist would respond by saying, “Nice job asking for a tickle,” and briefly tickled the student. If a student manded an unavailable item, the therapist would say, “Nice job asking for X, maybe we can have some soon.” RIRD resembled baseline, in that verbal praise was delivered for appropriate verbal responses. However, occurrences of vocal stereotypy were interrupted and then redirected to other vocalizations, such as social questions, or vocal imitation. Results showed the RIRD successfully decreased rates of stereotypy for all participants.

### **Multiple Schedule**

Ahearn, Clark, Gardenier, Chung, and Dube (2003) used a multiple schedule design to evaluate the effects of external reinforcers on the persistence of automatically maintained behavior. The study included three participants who engaged in either vocal or motor stereotypy. Following a competing items assessment, in which two items which were preferred and competed with automatic reinforcers were identified, participants were exposed to two alternating multiple schedule sequences, each consisting of four

components. The first sequence was a behavioral momentum sequences (B-MO) consisting of baseline (no preferred stimulus available), variable-time (VT) exposure (first preferred stimulus available), test (second preferred stimulus available), baseline. The second sequence was a control sequence. During control, the sequence was baseline, baseline, test, and baseline. Results indicated that persistence of stereotypy increased following exposure to external reinforcers.

In 2006, Rapp evaluated the effects of noncontingent matched stimulation (NMS) and response blocking on stereotypic behavior using a multiple schedule that consisted of three 15-min components. There was one participant in this study, a 9-year-old boy who had been diagnosed with autism and mental retardation. The target behavior was object tapping which was defined as contact of a finger or hand to a surface (e.g., table, chair) with simultaneous movement that generated an audible product. Components of the multiple schedule were preintervention, intervention, and posintervention. This three-component sequence was conducted three times for both NMS and response blocking.

During the preintervention component, the participant was seated at a table that was devoid of materials or toys. A trainer sat by the door and provided no social consequences for his behavior. During the intervention component, either NMS or response blocking was implemented. When NMS was implemented, the participant was provided continuous access to four matched toys that had been identified during a preference assessment. When response blocking was implemented, the trainer sat by the participant, rather than by the door, and no toys were present. The trainer placed, but did not hold, the participant's hands in his lap contingent on object tapping. The postintervention component was the same as preintervention, and evaluated the extent to

which stereotypy increased (relative to preintervention levels) following intervention with NMS and response blocking. The presence of toys during NMS, and the location of the trainer during response blocking, signaled the contingency of the intervention component in effect. Results showed that object tapping always occurred at higher levels after response blocking relative to before response blocking. Also, the behavior always occurred at lower levels after NMS than before NMS. These results suggest that response blocking may have produced deprivation for the product of stereotypy. In addition, NMS may have provided stimulation that was similar to the product of stereotypy.

In 2007, Rapp further evaluated the use of the multiple schedule as a method to identify matched stimulation that may substitute for stimulation accessed through engaging in automatically reinforced vocal stereotypy. Results of this study showed that the vocal stereotypy of both participants occurred at lower levels than the preintervention component, following exposure to auditory stimulation.

### **Purpose**

The use of a brief functional analysis has been shown to be effective in evaluating behaviors that occur consistently and at moderate to high levels (Derby et al., 1992). In addition, the use of brief functional analyses and preference assessments have the ability to aid the clinician in moving more efficiently from the assessment of problem behavior into treatment. Use of a multiple schedule sequence, has been shown to be an effective method in manipulating levels of stereotypic behavior that have been shown to occur consistently at moderate to high levels (Ahearn et al., 2003; Rapp, 2006, 2007). Due to the structure of the multiple schedule in which several components can be alternated

within the same sequence in one session, it serves as possible brief method to rapidly identify treatment effects.

The primary purpose of this study was to evaluate the use of a multiple-schedule design as a method for rapidly identifying treatment effects. A secondary purpose of this study was to extend the findings of Ahearn et al. (2007) by determining the relative efficacy of motor versus vocal RIRD for decreasing rates of vocal stereotypy.

## **Method**

### **Participants**

Ben was an 8-year-old boy who had been diagnosed with an autism spectrum disorder and received clinical and educational services in a day services setting. Ben communicated vocally to request needed or desired items and activities, imitate, respond to greetings, answer a variety of social questions, as well as participate in various academic activities throughout his day. Ken was an 11-year-old boy who had been diagnosed with an autism spectrum disorder and received clinical and educational services in a day services setting; he lived with his parents. Ken was able to communicate vocally in order to request needed or desired items and activities, imitate, respond to greetings, answer a variety of social questions, as well as participate in various academic activities throughout his day. Both participants' vocal stereotypy consisted primarily of delayed echolalia and repetitive speech and sounds

Ben and Ken were included in this study because both educational and clinical service providers noted that their vocal stereotypy interfered with educational services and persisted across environments and activities.

### **Setting and Materials**

Functional analysis sessions were conducted in a room (1.5 m by 3 m) equipped with wide-angle video camera, microphone, video recording equipment, materials necessary to conduct the experimental conditions, and a table with two chairs. Due to interfering behaviors that occurred while in this room, Ben's functional analysis was conducted in a cubicle in his classroom, where he typically received instructional programming throughout his school day.

All treatment assessment sessions were conducted in classroom cubicles and were recorded using a hand held mini digital video recorder placed on a tripod or flat surface.

### **Response Measurement and Interobserver Agreement**

*Vocal stereotypy* was defined as any instance of non-contextual speech, including singing, babbling, and repetitive production of actual words. This definition did not include speech paired with eye contact to another person, requesting items, words produced in response to teacher directives, or recognizable words produced in the context of play. During functional analysis sessions, data on vocal stereotypy were collected using 10-s momentary time sampling. During treatment sessions, Vocal stereotypy data were collected using continuous duration recording, and were converted into a percentage of occurrence per session by dividing the total number of seconds of vocal stereotypy by the total number of seconds in the session and multiplying by 100.

Two independent observers recorded responding for at least 30% of sessions across all phases, conditions, and participants. For the functional analysis, interobserver agreement was calculated by dividing the number of interval agreements by the number of agreements plus disagreements and multiplying by 100. Mean total agreement for vocal stereotypy was 90% (range, 77%-95%) for Ben and 95% (range, 87%-100%) for



Ken. For treatment phases, exact agreement (total seconds of vocal stereotypy in a session recorded by each observer) was calculated and mean agreement was 93% (range, 79%-98%) for Ben, and 91% (range, 70%-100%) for Ken for the brief treatment analysis, and 95% (range, 87%-100%) for Ben, and 97% (range, 93%-100%) for Ken for the extended treatment analysis.

### **Functional Analysis**

During the functional analysis of vocal stereotypy, alone, attention, and demand conditions were alternated using a multi-element design. The sequence of conditions was alone, alone attention, alone, alone, demand (Roscoe, Carreau, MacDonald, & Pence, 2008). Sessions lasted for 5 min.

During the alone condition, the participant was left alone in a room equipped only with two chairs and a table, and was monitored using a two-way mirror. For Ben, the alone condition was modified to a no-interaction condition in which the teacher remained in the room. The teacher would begin a no-interaction session by stating, “Ben, I’m just going to be over here doing some work” and looked at a magazine or otherwise appeared busy. This modification was made in order to keep Ben from engaging in dangerous behavior that had previously occurred during the alone condition.

During the attention condition, the therapist presented moderately preferred leisure activities, identified from a recent preference assessment, to the participant and delivered verbal attention (e.g., “Ben, stop that.”) contingent on the occurrence of vocal stereotypy.

During the demand condition, the therapist presented academic demands which were removed for 15 s contingent on the occurrence of vocal stereotypy. Demands were

those typically seen in the participant's regular academic programming and were not mastered (80% or less performance accuracy and independence). Results for the functional analysis showed that vocal stereotypy was automatically maintained for each participant (Figures 1 and 2).

### **Procedure**

**Baseline.** During baseline sessions, a moderately preferred item was presented to the participant. When the participant emitted vocal stereotypy, the therapist neutrally removed the item. The therapist redelivered the item contingent on participant mands (i.e. "I want X.") for the item in the absence of vocal stereotypy. This condition approximated contingencies that typically occurred in the natural environment. If the participant manded for an item other than the moderately preferred item presented at the start of the session, the therapist would respond, "nice asking, you can have X later." If the participant tacted (i.e., "That's an X.") an item in the room the teacher would respond, "that's right, that's a(n) X." All other behaviors were ignored.

**Response Interruption and Redirection-motor (RIRDm).** During response interruption and redirection-motor (RIRDm) sessions, a moderately preferred item was present, but out of the student's reach. Least to most prompting was used to prompt the participant to mand for the item. If the participant manded for the item it was delivered by the therapist. When the participant emitted vocal stereotypy, the therapist neutrally removed the item, then prompted known motor responses (e.g., touch nose, touch mouth, clap hands, etc.) until the participant independently complied with 3 consecutive motor responses in the absence of vocal stereotypy. Immediately following the procedure, the therapist again prompted the participant to mand for the item. In order to allow the

participant equal opportunities to engage in vocal stereotypy across baseline and treatment sessions, the session timer was stopped at the onset of the RIRDm procedure and was started following the participant independently complying with 3 consecutive motor responses. All other conditions remained the same from baseline.

**Response Interruption and Redirection-vocal (RIRDv).** Response interruption and redirection-vocal (RIRDv) sessions were conducted in the same manner as RIRDm conditions with one modification: following vocal stereotypy the teacher would present vocal compliances (e.g., “What’s your name?”, “Where do you go to school?”) until the participant independently complied with 3 consecutive vocal responses in the absence of vocal stereotypy.

### **Experimental Design**

**Brief Treatment Analysis.** A multiple schedule was used in order to rapidly detect the effect of different conditions on the rate of behavior. Multiple schedule sessions were conducted once per day and lasted for 15 min. Sessions were composed of a 5-component sequence which alternated baseline and treatment conditions. Each sequence component was 3 min. One multiple schedule sequence was conducted each day, across 6 days.

There were two sequences run. The first was an RIRDm sequence which consisted of baseline, RIRDm, baseline, RIRDm, and baseline. The second was an RIRDv sequence, which consisted of baseline, RIRDv, baseline, RIRDv, baseline. Three RIRDm sequences were run followed by three RIRDv sequences.

**Extended Treatment Analysis.** An ABABACAC design in which baseline (A), RIRDm (B), and RIRDv (C) were alternated to determine the effects of the treatment

procedures on vocal stereotypy. Between 2 and 5 sessions were conducted each day, each lasting 5 min.

### Results

Results of Ben's brief treatment analyses can be seen in Figures 3 and 4. Figure 3 shows results from the series of RIRDm component sequences. During the initial baseline of component sequence 1, vocal stereotypy occurred at a high level, and dropped to a moderate level in the first RIRDm treatment component. During the second baseline component, vocal stereotypy also occurred at moderate but slightly lower level than that seen in the first treatment component. The level of vocal stereotypy continued to decrease during the second treatment component, but recovered slightly during the final baseline of the first sequence. During the second RIRDm component sequence, vocal stereotypy occurred at a moderate level during the first baseline, and decreased to a low level in the first RIRDm treatment component. During the second baseline component, vocal stereotypy recovered to a moderate level, followed by a modest decrease in the second RIRDm component, and only a very slight increase in the level of vocal stereotypy was seen in the final baseline of the second sequence. However, baseline levels never fell below those seen in the treatment component during the second sequence. During the final component sequence of the RIRDm brief treatment analysis, vocal stereotypy occurred at a moderate level during the first baseline then decreased to a low level in the first treatment component. The level of behavior continued to decrease in the second baseline, and showed a slight increase in the second RIRDm treatment component,

followed by a recovery to a moderate level in the final baseline. A proportional analysis of these data can be found in Figure 5.

Proportional data were calculated by dividing the averaged percent occurrence of vocal stereotypy during treatment by the average percent occurrence of vocal stereotypy during baseline for each component sequence. If proportions were equal to, or greater than one, this would indicate that no treatment effect was obtained. While visual representation of the data for Ben's RIRDm sequences shows an overall decreasing trend across sequences, these proportional data show that the treatment effect remained stable across all three sequences. Stereotypy occurred 0.67, 0.69, and 0.63 ( $M=0.67$ ) the levels seen in baseline across the three RIRDm component sequences.

Results of Ben's RIRDv brief treatment analysis can be seen in Figure 4. In the first RIRDv component sequence, both treatment components show significantly lower levels of vocal stereotypy than those that occurred during baseline. With the exception of the first baseline in the third sequence, levels of stereotypy remain high during the baseline components across all three sequences, and all treatment components show significant decreases to low or moderate levels. Mean baseline level was 64% in baseline and 25% during treatment across all three sequences. A proportional analysis of these data can be found in Figure 6. Similar to the RIRDm series, treatment effect was stable across all three sequences, where vocal stereotypy occurred at 0.40 the levels seen in baseline in each sequence.

Figure 7 shows the results of Ken's RIRDm brief treatment analysis. During the initial baseline component of the first sequence vocal stereotypy occurred at high level followed by a modest decrease in the first treatment component. Vocal stereotypy

recovered to a high level in the second baseline component, followed by a modest decrease in the second RIRDm treatment component, and no change in the final baseline. During the first and last baseline components of the second sequence, vocal stereotypy occurred at high levels. Stereotypy occurred at low levels in both treatment components as well as the second baseline. During the final RIRDm sequence, vocal stereotypy occurred at a higher level than treatment only in the second baseline. Proportional data for Ken's RIRDm brief treatment analysis are in Figure 9. The treatment effect across sequences was variable with treatment levels of vocal stereotypy occurring at 0.51, 0.25, and 0.68 ( $M=0.45$ ) the levels seen in baseline.

During the first sequence of Ken's RIRDv brief treatment analysis (Figure 8), vocal stereotypy occurred at a level equal to or less than the levels observed during treatment. During the second baseline of the second RIRDv sequence, the level of vocal stereotypy was similarly equal to or less than levels seen in the treatment components of that sequence. During the final baseline of the second sequence, vocal stereotypy occurred at a lower level than that seen in the first treatment component of that sequence, but showed an increase in level from the second treatment component. During Ken's third RIRDv component sequence, the highest level of stereotypy occurred during the second treatment component. Proportional data for Ken's RIRDv brief treatment analysis are in Figure 10. Again, the treatment effect across sequences was variable with treatment levels of vocal stereotypy occurring 0.33, 0.88, and 0.75 ( $M=0.73$ ) those seen in baseline.

Results of Ben's extended treatment analysis are in Figure 11. During the initial baseline vocal stereotypy occurred with an increasing trend from moderate to high levels, followed by an immediate decrease to low levels when RIRDm was introduced. During

the return to baseline, levels initially remained low in the first session, but recovered to moderate levels, and then decreased back to low levels during the reversal to RIRDm. During the second return to baseline, levels were initially low to moderate and variable, but recovered moderate rates similar to those in the second baseline. When RIRDv was introduced, levels of vocal stereotypy did not immediately decrease, but showed an variable decreasing trend from moderate to low levels, and stabilized at low levels in the final three sessions. During the final return to baseline, levels of vocal stereotypy immediately recovered to moderate to high levels, and immediately decreased when RIRDv was reintroduced. Table 1 shows proportional analysis data for Ben's extended analysis. The most profound treatment effect was seen in the return to RIRDv from the final baseline where levels were 0.29 those seen in the preceding baseline. However, treatment effect of RIRDm was stronger than that seen in the brief analysis (0.52, 0.43). When comparing average treatment levels for both RIRDm and RIRDv to average levels of vocal stereotypy seen across the four baseline conditions, the behavior occurred at 0.48 during RIRDm, and 0.41 in RIRDv.

During the initial extended baseline condition (Figure 12), Ken's vocal stereotypy never recovered to levels comparable to the highest levels seen during his brief treatment analysis, therefore therapists were unable to introduce treatment.

### **Discussion**

Treatment effects for Ben remained consistent across component sequences for both the RIRDm and RIRDv brief treatment analyses, while results were variable for Ken. Ben's extended treatment showed the greatest treatment effect in the final presentation of RIRDv, although the effects of the treatment were generally equal across

both interventions. Ken's extended analysis was unable to be completed due to low levels of vocal stereotypy in the initial baseline condition. This could be contributed to carry-over effects from the multiple schedule or, given the inconsistency of results obtained from his brief treatment analysis, his behavior may have been controlled by variables other than the contingencies in place during experimental conditions.

During Ben's brief treatment analysis, treatment effects were strongest during the RIRDv component sequences. Similarly, the strongest effect was seen when RIRDv was introduced during the second time in Ben's second analysis. Also the effect of RIRDm was stronger in the extended analysis (0.52, 0.43) than during the brief analysis ( $M=0.67$ ). When comparing the average of both treatments divided by the average of all four baseline conditions, it is found that both RIRDm and RIRDv had a similar effect on Ben's vocal stereotypy during the extended analysis. While there is some variability across the brief and extended analyses for Ben in regards to treatment effect, the treatment effects seen in the both the RIRDm and RIRDv brief analyses, were validated by the treatment effects observed in the extended analysis. This lends some support to the use of a multiple schedule as a reliable method for evaluating treatment effects. However, due to the inability to complete the extended analysis on Ken, it is unclear whether or not this method will generalize across participants. If this method can be shown to be an effective method for evaluating treatment effects across participants, future studies should evaluate its validity across a variety of problem behaviors and treatments.

This study is an extension of previous literature in two ways. First, it replicates the findings of Ahearn et al. (2007), in that one participant's vocal stereotypy was



decreased using response interruption and redirection. It also supports the use of the brief treatment analysis as a means to identify treatment effects (Rapp, 2006; 2007). However, as previously stated, further analysis is needed to determine whether the brief treatment analysis is a reliable way in which to identify treatment effects across participants.

### References

- Ahearn, W. H., Clark, K. M., & Gardenier, N. C. (2003). Persistence of stereotypic behavior: examining the effects of external reinforcers. *Journal of Applied Behavior Analysis, 36*, 439-448.
- Ahearn, W. H., Clark, K. M., & MacDonald R. P. F. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis, 40*, 263-275.
- Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalist context. *Journal of Applied Behavior Analysis, 33*, 353-357.
- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519-533.
- Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: a summary of 79 cases. *Journal of Applied Behavior Analysis, 25*, 713-721.
- Fellner, D. J., Laroche, M., & Sulzer-Azaroff, B. (1984). The effects of adding interruption to differential reinforcement on targeted and novel self-stimulatory behaviors. *Journal of Behavior Therapy and Experimental Psychiatry, 15*, 313-321.
- Fisher, W., Piazza, C. C., & Bowman, L. G. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal*

- of Applied Behavior Analysis*, 25, 491-498.
- Hagopian, L. P., & Adelinis, J. D. (2001). Response blocking with and without redirection for the treatment of pica. *Journal of Applied Behavior Analysis*, 34, 527-530.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994/1982). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, 27, 197-209.
- Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., & Smith, R. G. (1994). The functions of self-injurious behavior: an experimental-epidemiological analysis. *Journal of Applied Behavior Analysis*, 27, 215-240.
- Lerman, D. C., & Iwata, B. A. (1996). A methodology for distinguishing between extinction and punishment effects associated with response blocking. *Journal of Applied Behavior Analysis*, 29, 231-233.
- Lewis, M. H., & Bodfish, J. W. (1998). Repetitive behavior disorders in autism. *Mental Retardation and Developmental Disabilities*, 4, 80-89.
- Lovaas, I., Newsom, C., & Hickman, C. (1987). Self stimulatory behavior and perceptual reinforcement. *Journal of Applied Behavior Analysis*, 20, 45-68.
- Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis*, 18, 249-255.
- Paramore, N. W., & Higbee, T. S. (2005). An evaluation of a brief multiple-stimulus preference assessment with adolescents with emotional-behavioral disorders in an educational setting. *Journal of Applied Behavior Analysis*, 38, 399-403.

- Rapp, J. T. (2006). Toward an empirical method for identifying matched stimulation for automatically reinforced behavior: a preliminary investigation. *Journal of Applied Behavior Analysis, 39*, 137-140.
- Rapp, J. T. (2007). Further evaluation of methods to identify matched stimuli. *Journal of Applied Behavior Analysis, 40*, 73-88.
- Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: A review of behavior assessment and treatment. *Research in Developmental Disabilities, 26*, 527-547.
- Rincover, A. (1978). Sensory extinction: a procedure for eliminating self-stimulatory behavior in developmentally disabled children. *Journal of Abnormal Child Psychology, 6*, 299-310.
- Roscoe, E. M., Carreau, A., MacDonald, J., & Pence, S. T. (2008). Further evaluation of leisure items in the attention condition of functional analyses. *Journal of Applied Behavior Analysis, 41*, 351-364.
- Vollmer, T. (1994). The concept of automatic reinforcement: implications for behavioral research in developmental disabilities. *Research in Developmental Disabilities, 15*, 187-207.
- Windsor, J., Piché, L. M., & Locke, P. A. (1994). Preference testing: a comparison of two presentation methods. *Research in developmental disabilities, 15*, 439-455.

Table 1

*Proportional Data Analysis from Ben's Extended Treatment Analysis*

Equation	Proportion	Equation	Proportion
RIRDm1/BL1	0.52	RIRDm1+2/BL1+2	0.48
RIRDm2/BL2	0.43	RIRDv1+2/BL3+4	0.41
RIRDv1/BL3	0.64	RIRDm1+2/BL1+2+3+4	0.48
RIRDv2/BL4	0.29	RIRDv1+2/BL1+2+3+4	0.47

## Figure Captions

*Figure 1.* Results of Ben's functional analysis. Vocal stereotypy occurred at highest levels during the no interaction condition, suggesting the behavior was automatically reinforced.

*Figure 2.* Results of Ken's functional analysis. Vocal stereotypy occurred at highest levels during the alone condition, suggesting the behavior was automatically reinforced.

*Figure 3.* Results of Ben's RIRDm brief treatment analysis.

*Figure 4.* Results of Ben's RIRDv brief treatment analysis.

*Figure 5.* Proportional data for Ben's RIRDm brief treatment analysis.

*Figure 6.* Proportional data for Ben's RIRDv brief treatment analysis.

*Figure 7.* Results for Ken's RIRDm brief treatment analysis.

*Figure 8.* Results for Ken's RIRDv brief treatment analysis.

*Figure 9.* Proportional data for Ken's RIRDm brief treatment analysis.

*Figure 10.* Proportional data for Ken's RIRDv brief treatment analysis.

*Figure 11.* Results for Ben's extended treatment analysis.

*Figure 12.* Initial baseline from Ken's extended treatment analysis.

Figure 1

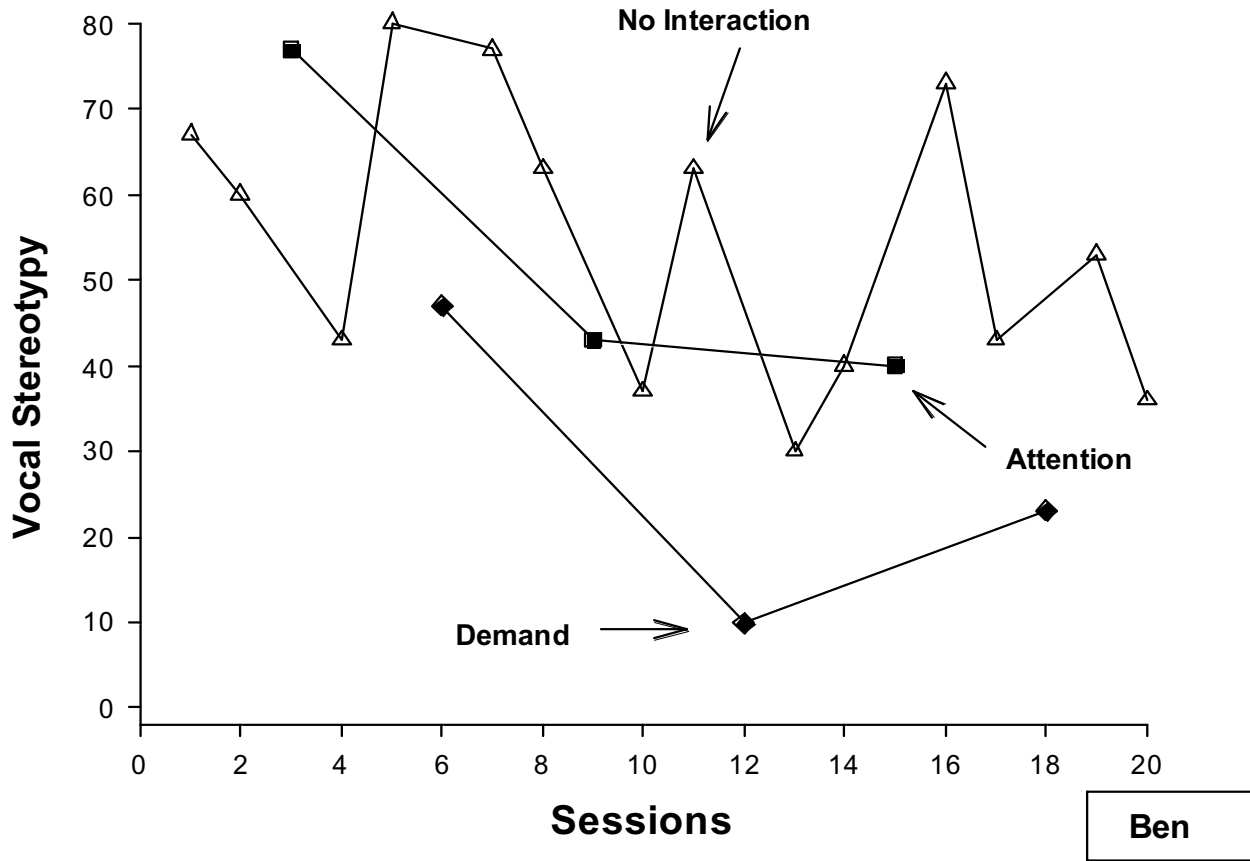


Figure 2

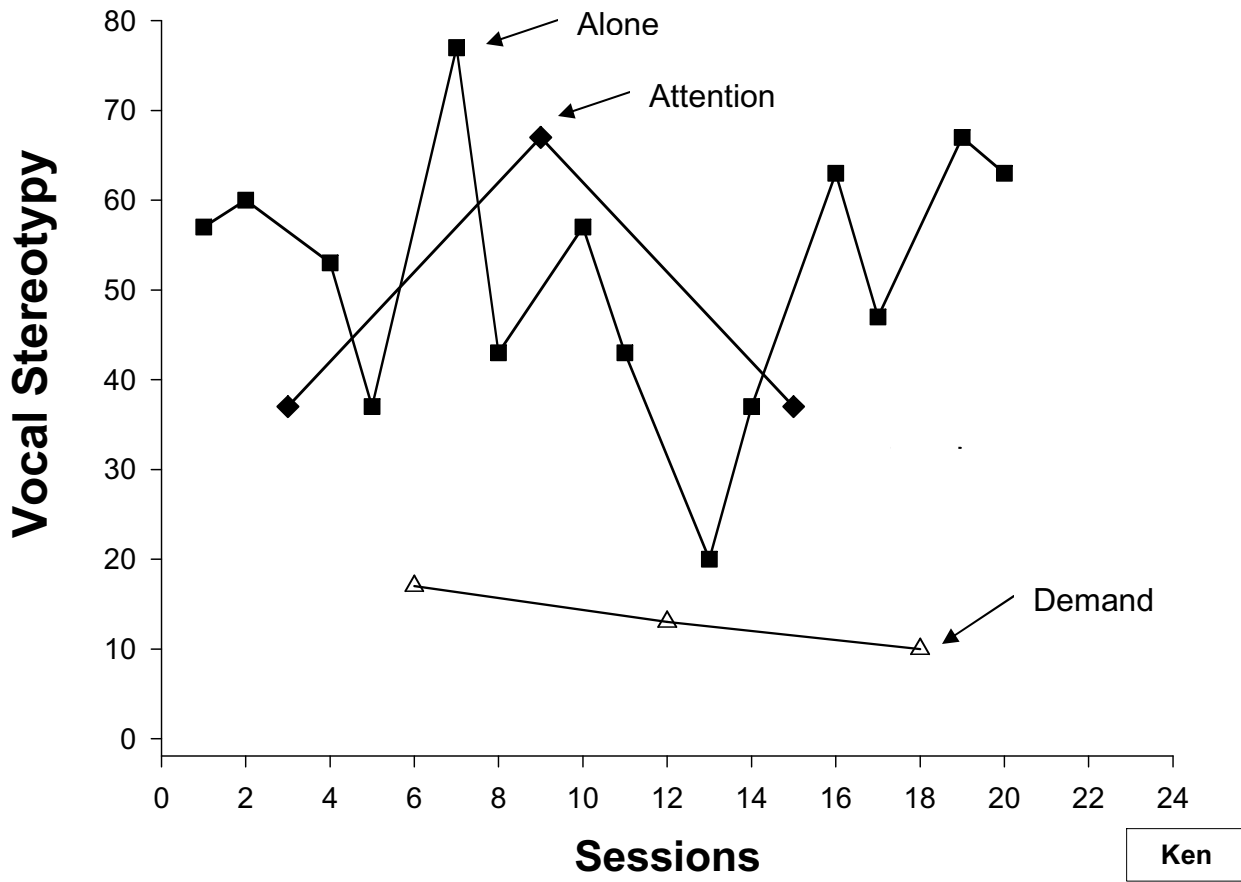




Figure 3

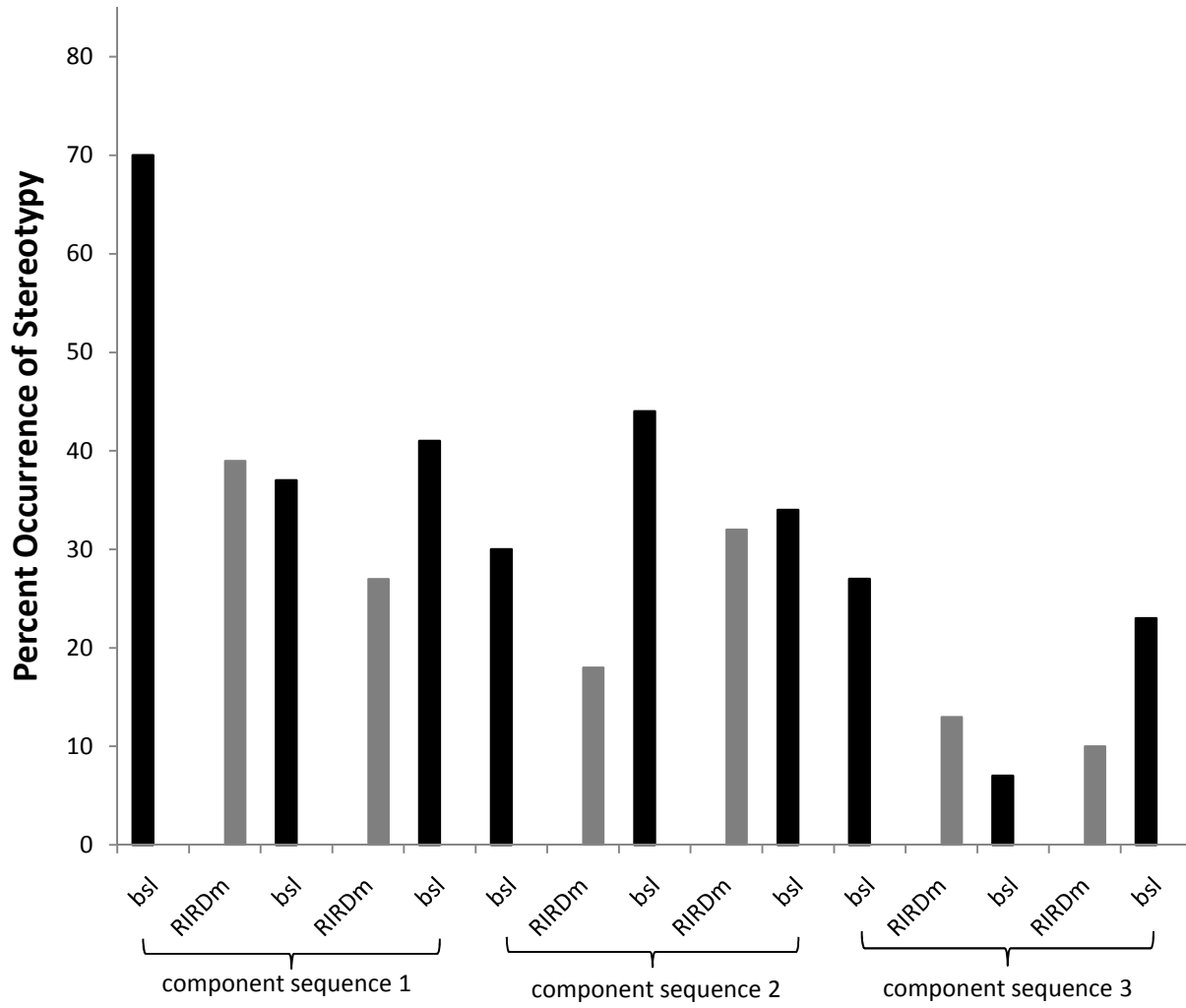


Figure 4

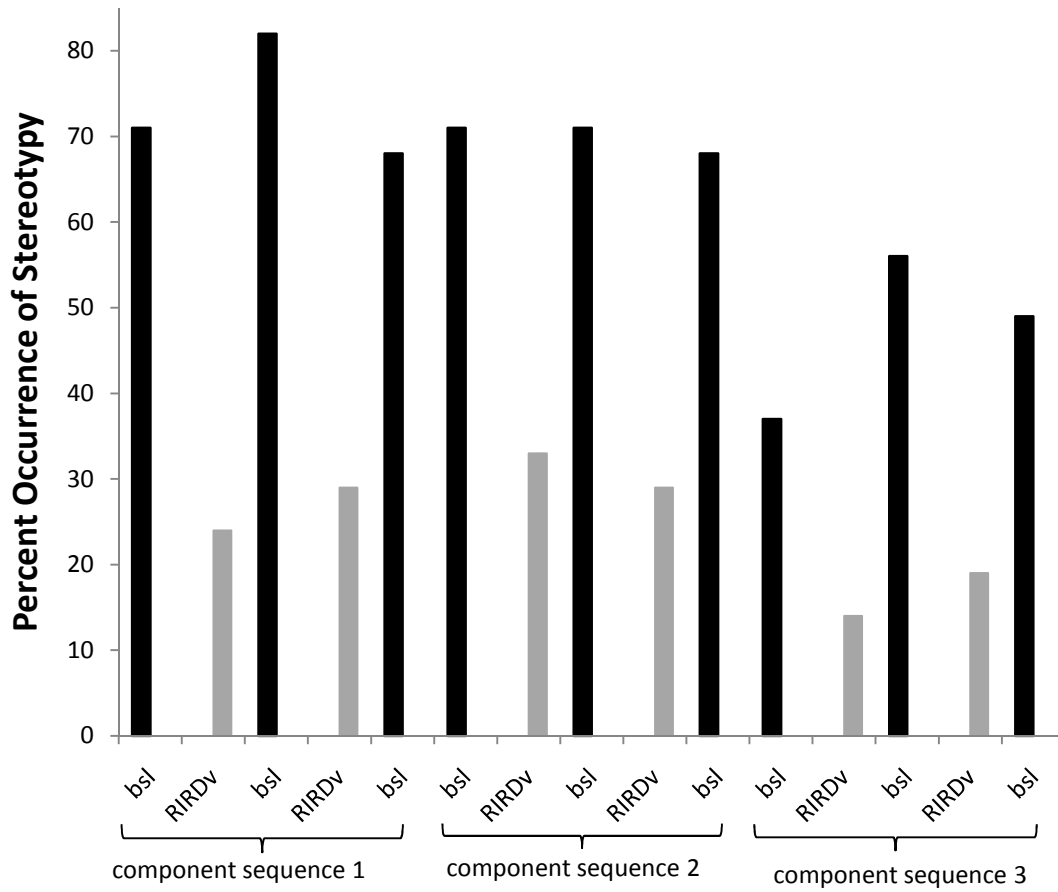


Figure 5

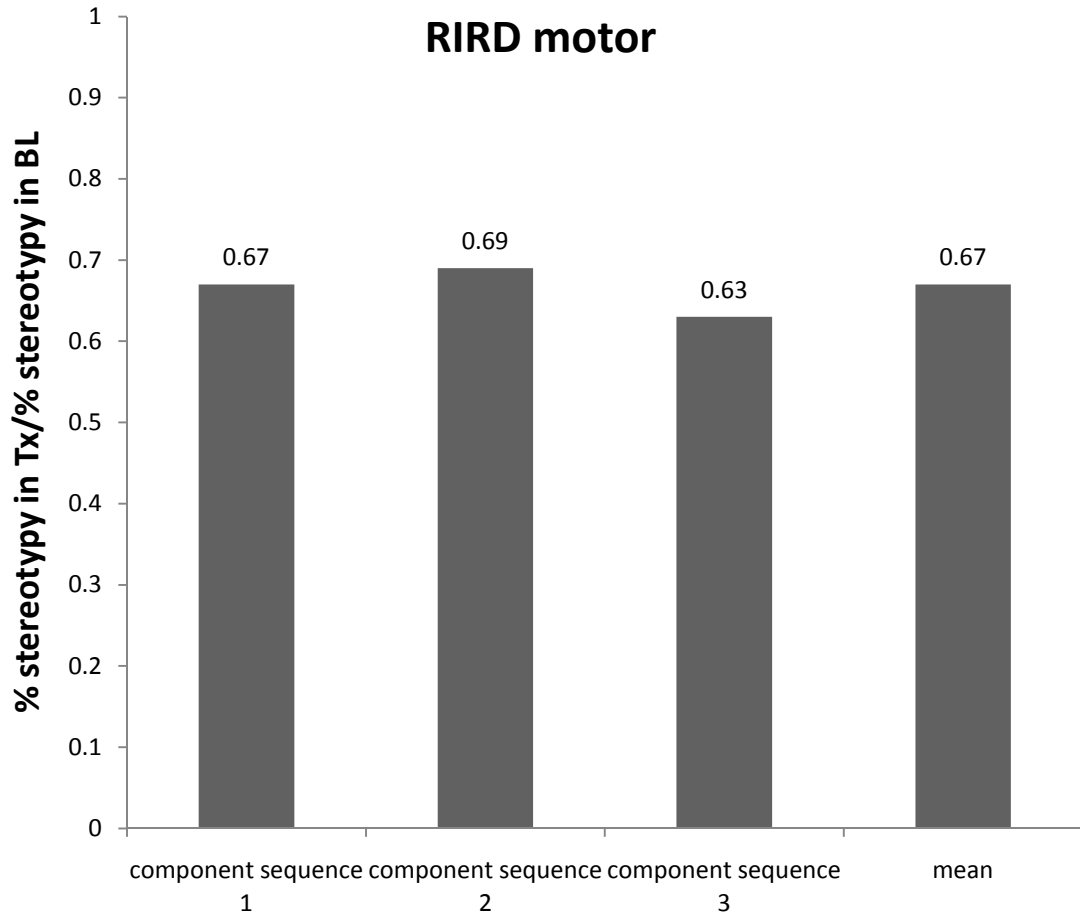


Figure 6

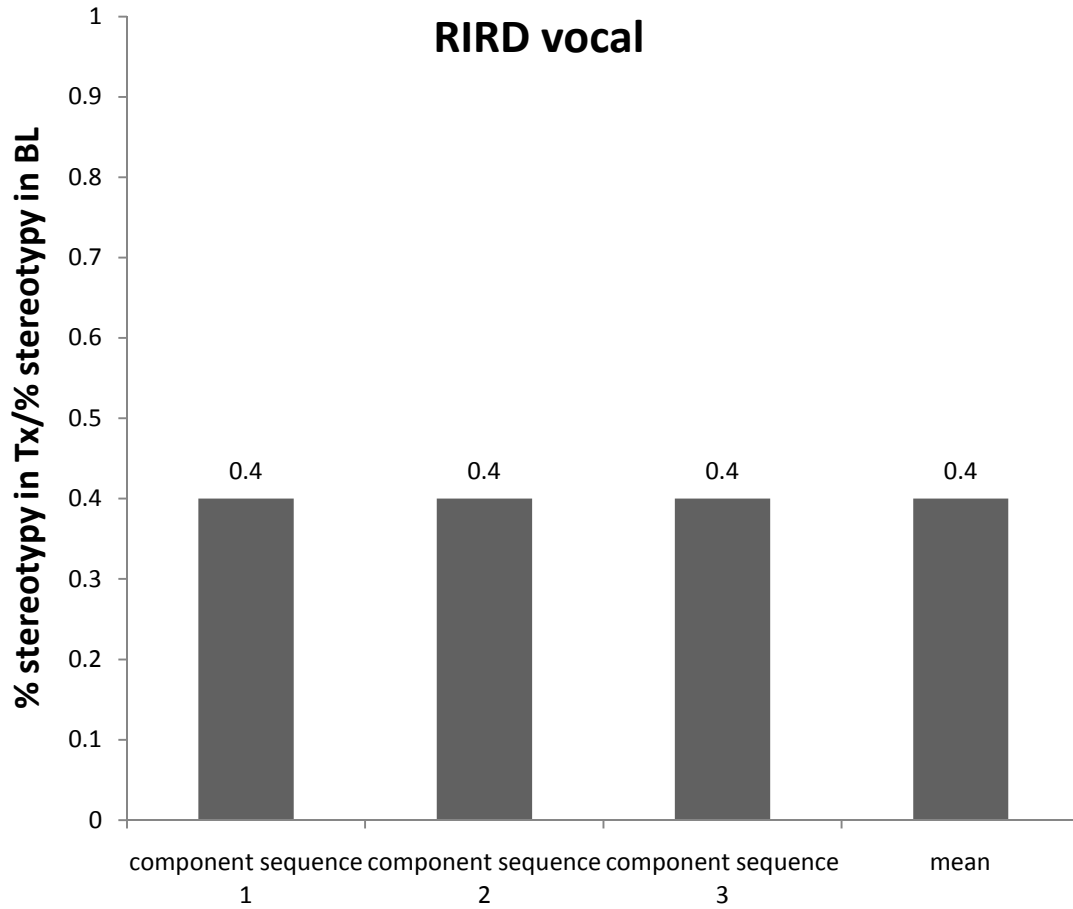


Figure 7

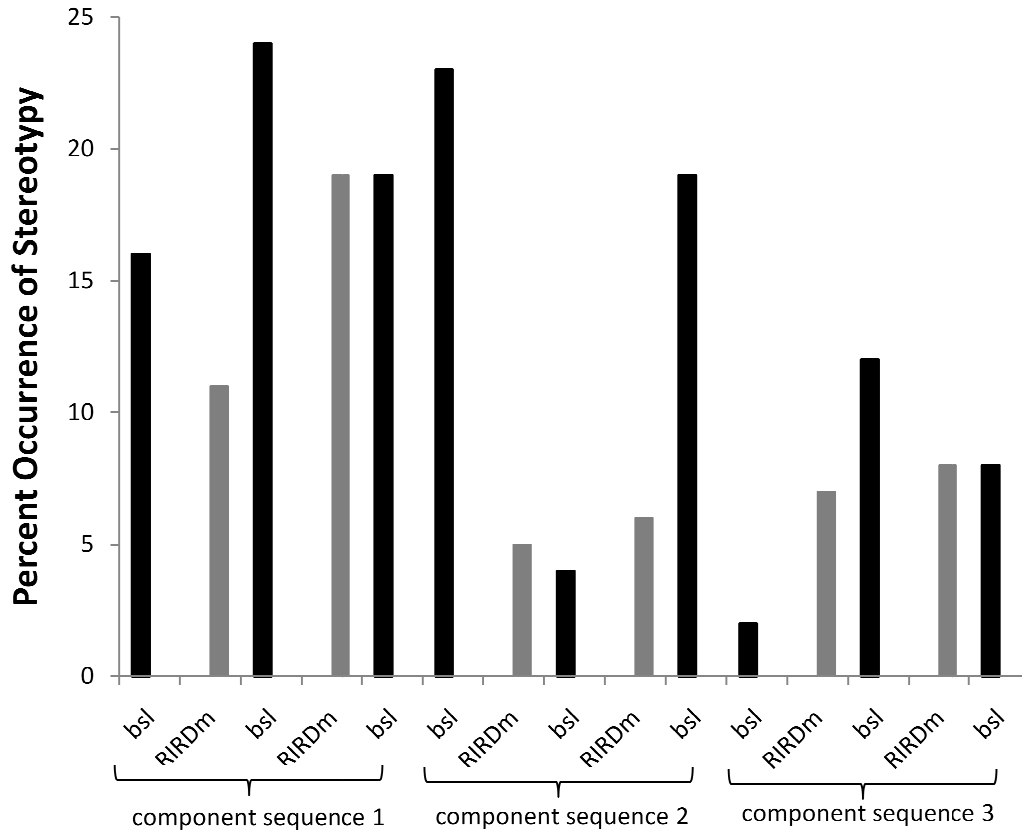


Figure 8

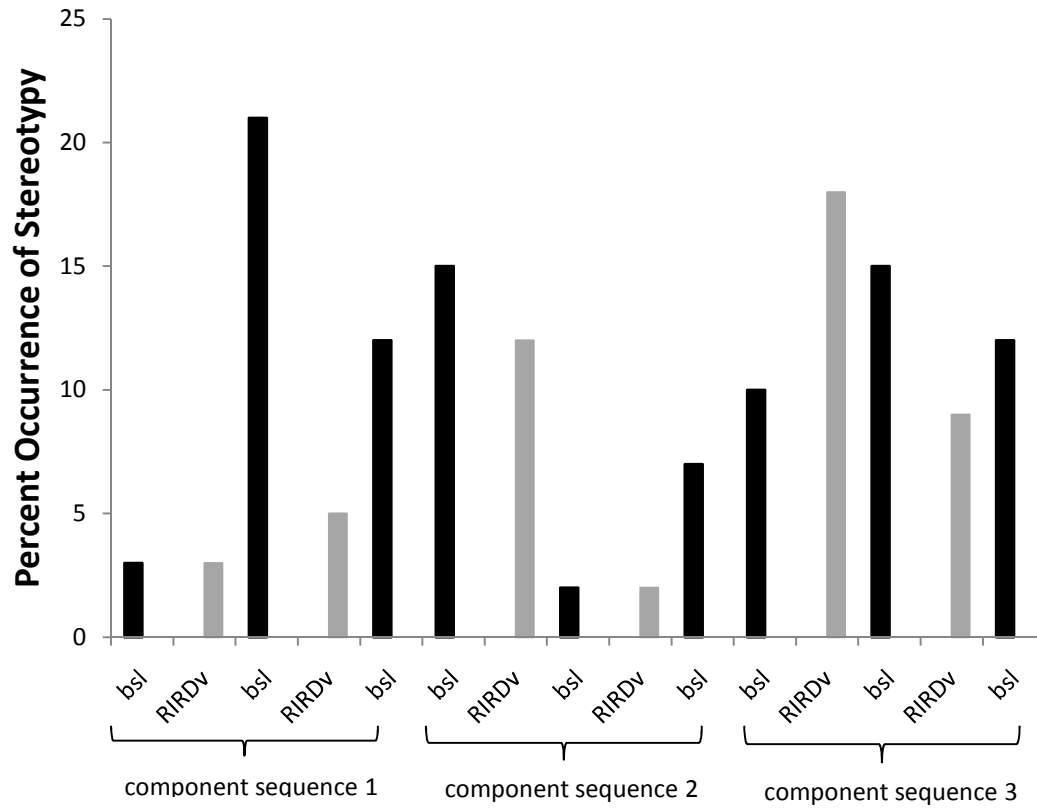


Figure 9

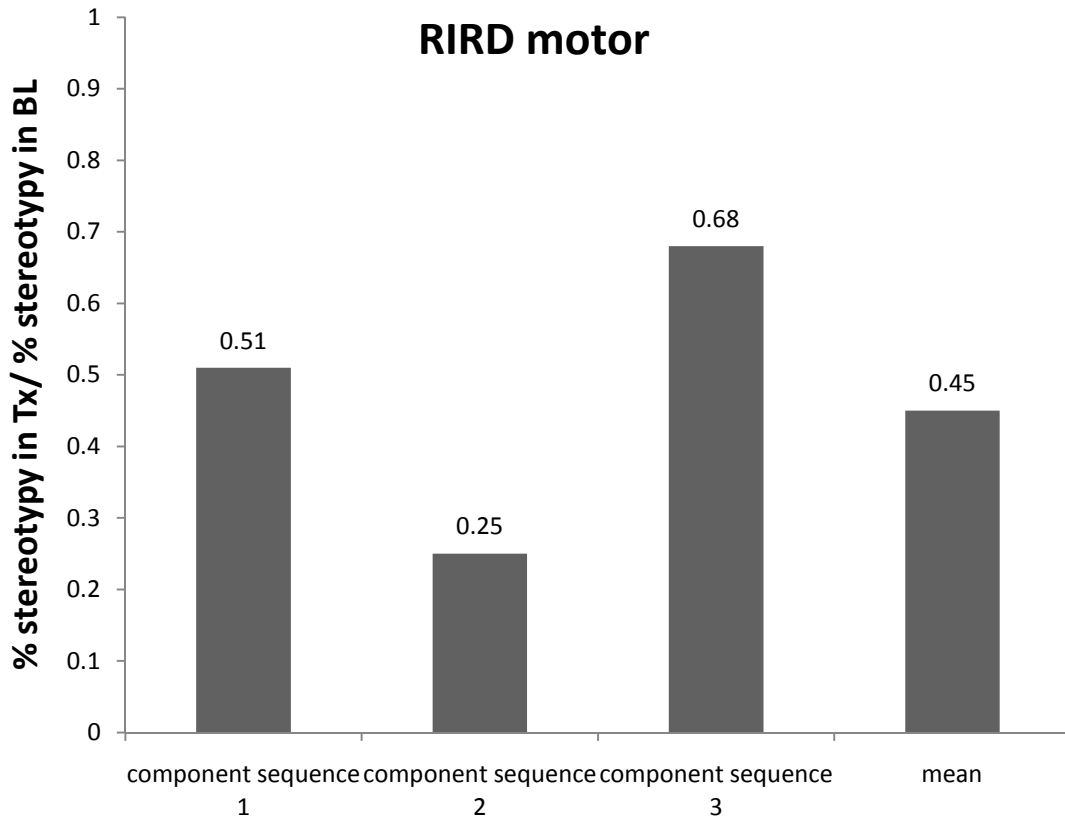


Figure 10

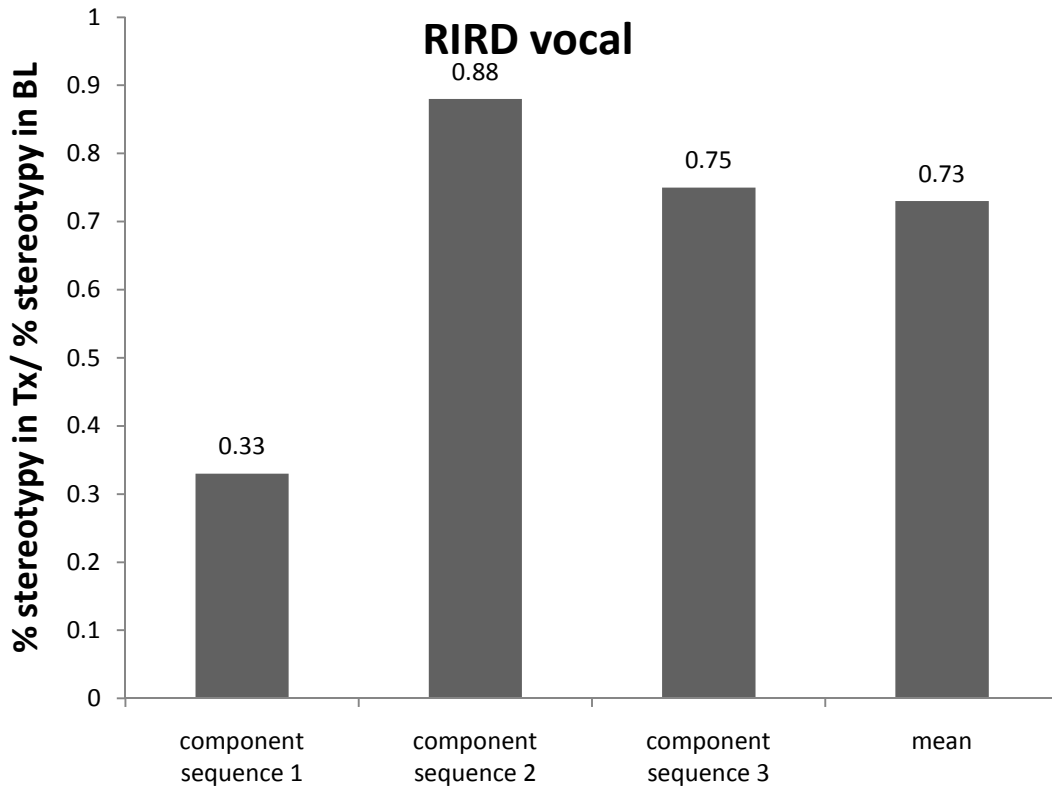




Figure 11

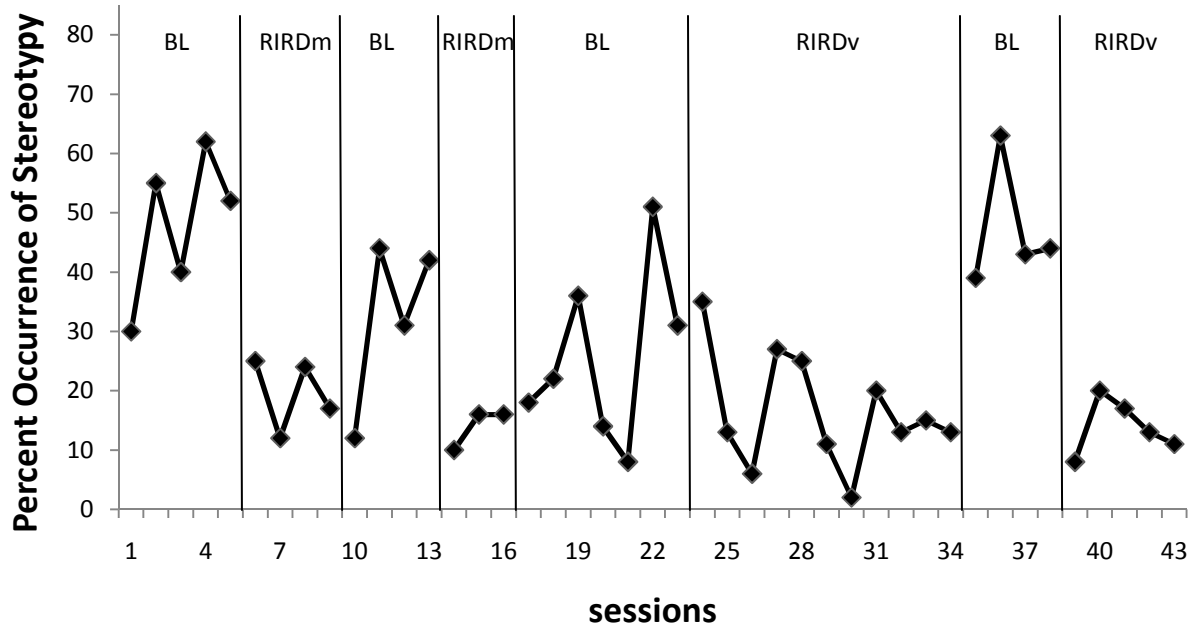


Figure 12

