7-2023

Further Evaluation of the Effects of Music and RIRD on Vocal Stereotypy

Keith Happel

Further Evaluation of the Effects of Music and RIRD on Vocal Stereotypy

by

Keith Happel

Bachelor of Arts

Applied Behavior Analysis

Florida Institute of Technology

2021

A thesis submitted to the School of Behavior Analysis

of Florida Institute of Technology in the partial fulfillment

of the requirements for the degree of

Master of Science

in

Applied Behavior Analysis

Melbourne, Florida

July, 2023

We the undersigned committee hereby approve the attached thesis, "Further Evaluation of the Effects of Music and RIRD on Vocal Stereotypy"

by

Keith Happel

_____

Kimberly N. Sloman, Ph.D., BCBA-D
Associate Professor
School of Behavior Analysis
Major Advisor

_____

Kaitlynn M. Gokey, Ph.D., BCBA-D
Assistant Professor
School of Behavior Analysis

_____

Renée N. Souris, Ph.D.
Assistant Professor
School of Arts and Communication

_____

Robert A. Taylor, Ph.D.
Professor and Dean
College of Psychology and Liberal Arts

Abstract

Title: Further Evaluation of the Effects of Music and RIRD on Vocal Stereotypy

Author: Keith Happel

Major Advisor: Kimberly Sloman

Vocal stereotypy (VS), or vocalizations that are noncontexual or nonfunctional, is commonly exhibited in individuals on the autism spectrum. The nature of vocal stereotypy may pose an issue for some individuals, ranging from social ostracization to the inability to perform daily tasks. Previous research has shown that access to matched stimulation (e.g., music) and response interruption and redirection (RIRD), a form of punishment, are effective at decreasing vocal stereotypy (Gibbs et al., 2018). However, in the aforementioned study, researchers did not evaluate a less intrusive treatment in isolation prior to implementing punishment. The purpose of the current investigation is to implement a least to most intrusive intervention for vocal stereotypy by evaluating matched stimulation first and then adding RIRD if necessary. We tested wearing headphones that played music as a form of matched stimulation to decrease VS and increase on-task duration. The data indicate that music as a form of matched stimulation significantly decreased VS for all participants across multiple settings. Additionally, marginal increases in on-task behavior were observed. To date, RIRD was not a necessary component for effective treatment.

Keywords: Autism Spectrum Disorder, Children, Matched Stimulation, RIRD, Vocal Stereotypy

Table of Contents

List of Figures

Chapter 1

History of Treatments for Vocal Stereotypy

Vocal stereotypy (VS) is a commonly observed response in individuals with autism spectrum disorder (ASD) and other developmental disorders. Vocal stereotypy can be defined as any instance of vocalizations that are noncontextual or nonfunctional (Ahearn et al., 2007). Some examples include babbling, repetitive phrases, squeals, and other vocalizations that are unrelated to their present environment. This behavior has been shown to interfere with acquiring adaptive behaviors. Sometimes individuals with vocal stereotypy are deemed "socially awkward" or "weird" due to the noncontextual nature of their speech (Ahearn et al., 2007). The deceleration of this behavior can lead to more learning opportunities for the individual including social skills and other communication skills.

Operant behavior is such a behavior that is freely emitted by the individual and can have its frequency manipulated through reinforcement or punishment. Operant behavior is also known as a Stimulus-Response-Stimulus relation. This three-term contingency is also described as discriminative stimulus, operant response, and reinforcer/punishment (Skinner, 1966). When looking at operant behaviors, two general classes of adaptive and maladaptive behaviors are studied. As described earlier, vocal stereotypy falls under the maladaptive side of operant behavior because it can interfere with the acquisition of other more socially valid behaviors.

Vocal Stereotypy

Vocal stereotypy is under the category of "self-stimulatory behavior," which is an operant behavior maintained by the automatic reinforcement produced by the behavior itself (Lovaas et al., 1987). Determining the function of this behavior is assessed through a functional analysis. The procedure for the most performed functional assessment (Iwata et al., 1994)

consists of the individual being exposed to multiple different conditions in which the parts of the three-term contingency are manipulated. The conditions are arranged to test if the target behavior is sensitive to different social reinforcers (e.g., attention from others, escape from demands) or if it persists in the absence of social consequences. Practitioners record the rate of target behavior across test (i.e., attention, escape, no consequence) and control (i.e., toy play) conditions and conduct each condition until clear patterns emerge. If responding is differentially higher in one or more test conditions relative to the control condition, then the function(s) have been identified. In Iwata's (1994) functional analysis, two-thirds of the participants had higher levels of SIB in a specific test condition. Since then, functional analyses have been used in a variety of situations and with a variety of modifications. These antecedent and consequent manipulations help determine the function of the behaviors in question.

Functional Analysis & Subtyping

The functional analysis has been criticized for multiple shortcomings such as the extended duration of the procedure and the theory that putting individuals into these manipulated situations could be dangerous for them and may increase the frequency of the aberrant behavior. These limitations were remedied when the brief functional assessment model was created (Northup et al., 1991). This model works best in situations in which the behavior is so severe that an extended phase would be detrimental to the individual's health, or when the therapist has limited time to run the assessment.

Querim and colleagues evaluated if a brief extended exposure to an alone or no-interaction session could be used as a screening procedure to determine if the target behavior may be automatically maintained (Querim et al., 2013). That is, response patterns during the extended ignore or alone condition were used to predict the function. If the behavior maintained

or increased in frequency during the alone condition, it was determined that the behavior was most likely automatically reinforced. If a decrease in frequency was found, the behavior was hypothesized as being maintained by social reinforcement. The screening results were compared to traditional functional analyses and results matched for 7 out of 8 participants. Querim's procedure may be especially relevant for vocal stereotypy as this behavior has been found to be most likely maintained by automatic reinforcement (Piazza et al., 2000; Rapp et al., 1999; Vollmer et al., 1994). The functional analysis model has been modified again recently with the addition of a subtyping model (Hagopian et al., 2015). This model was originally created to categorize automatically reinforced SIB in terms of its responsiveness to environmental stimulation.

Three subtypes were created based on the patterns of responding recorded during a functional analysis and if the individual showed signs of self-restraint. Subtype 1 was scored if the automatically reinforced SIB had a quotient score greater than or equal to 0.5. This means that there was a clear differentiation in the levels of SIB between the alone and play conditions, with the play condition having significantly lower rates. This result suggests that access to a preferred reinforcer such as a toy may be competing with the SIB. Subtype 2 is chosen if the participant has a quotient score less than 0.5, has 30% or more overlapping data points between the two conditions, or if the mean rate of SIB was higher than 50 responses per minute in both conditions. This subtype is chosen for individuals who show high levels of SIB regardless of conditions. Subtype 3 is recorded when self-restraint occurs at least 25% of the trials in at least 3 consecutive alone conditions, and that the SIB was maintained by automatic reinforcement after the self-restraint was blocked. In this subtype, it is suggested that the self-restraint by the individual is competing with the SIB. These subtypes were adapted and used in a study where

the validity of reinforcement-based interventions on these subtypes were tested (Wunderlich et al., 2022). Through a literature search of automatically maintained stereotypy, the researchers determined which subtype the targeted stereotypy would fall into, and what interventions were effective or ineffective.

Unfortunately, there was not a single intervention that was effective for every participant, but that is to be expected with the idiosyncratic nature of behavior. This research does further the idea that, for Subtype 1 automatically reinforced behaviors, access to items may be competing with the behavior, and that Subtype 2 may require additional consequence-based strategies in most cases. As shown in the data presented by Wunderlich, interventions such as noncontingent reinforcement were more effective on Subtype 1 than on Subtype 2. Furthermore, while 62% of the included Subtype 1 studies achieved positive outcomes, only 49% of the Subtype 2 studies had successful results.

Matched Stimulation

Research on inappropriate vocal behavior such as echolalia and vocal stereotypy have found multiple potential interventions to decrease this stereotypy. One of these interventions is an intervention known as matched stimulation (MS; Piazza et al., 2000). Piazza et al. (2000) investigated the use of MS on different forms of aberrant behavior such as mouthing and dangerous climbing. In their study, they conducted a standard functional analysis (Iwata et al., 1982/1994) to determine the function of the behaviors. Participants then experienced a stimulus preference assessment using the procedures described by (Piazza et al.,1996). The items chosen were hypothesized to match the stimulation provided by the aberrant behaviors, and items that would provide a different kinesthetic consequence. They found that the stimuli that were hypothesized to match the same consequences resulted in lower levels of the target behavior.

While this study was conducted on behaviors such as mouthing, the effects also have implications for stereotypic behaviors. Their research found that when items are given to the individual that match the hypothesized consequence of the aberrant behavior, it is more likely to decrease the behavior rather than using items from a preference assessment or at random.

The results from Piazza et al. highlight the need for individualized assessment for stimuli used in interventions. An assessment that is commonly used in these studies is the competing items assessment. During this procedure, researchers provided the individuals with free access to preferred stimuli one at a time for a brief amount of time. Problem behavior data are recorded during these sessions as well as the duration of time spent interacting with the stimuli. Stimuli with increased engagement and decreased problem behavior are selected as competing items. These assessments are helpful in finding components that can be used as a part of an intervention for automatically reinforced problem behaviors (Piazza et al., 2000).

The history of MS involving vocal stereotypy and musical interventions involving non-contingent reinforcement begins with SIB, another automatically reinforced behavior. Some of the earliest literature of competing stimuli and matched stimuli decreasing automatically reinforced behaviors come from a study in which pica and three other potentially dangerous target behaviors of different functions were decreased when an intervention was put in place that competed with the function of those behaviors (Piazza et al., 2000). Matched stimulation was introduced as an intervention in the form of auditory stimulation for vocal stereotypy. In their study, researchers compared fixed-time delivery of an auditory toy to a condition in which the toy was delivered on a differential reinforcement of other behavior (DRO) schedule (i.e., contingent upon the absence of vocal stereotypy; Taylor et al., 2005). During pre-assessments, free access to auditory toys reduced vocal stereotypy relative to non-auditory toys. However,

access to the toys on the fixed-time schedule was ineffective at reducing stereotypy and a DRO schedule was necessary for treatment effects.

Other researchers have evaluated the effects of continuous access to music and other auditory stimuli on vocal stereotypy. One group of researchers tested the effects of multiple different types of auditory stimuli on vocal stereotypy. Researchers used noncontingent access to music, a white noise recording, and recordings of the participants' own vocal stereotypy (Saylor et al., 2012). White noise can be described as a sound that contains every frequency within the range of human hearing. Saylor tested these auditory stimuli in a quasirandom order of white noise, music, and their own vocal recordings. While white noise did not seem to decrease the rate of vocal stereotypy, both music and recordings of their own stereotypy lowered the rates to nearly non-zero levels. While the self-recordings had a similar efficacy of lowering vocal stereotypy, noncontingent music had higher social validity scores and was more preferred by the caretakers and participants.

Researchers also investigated the effects of noncontingent music on vocal stereotypy, however they focused on if the music itself was highly or lowly preferred and then recording the rate of the behavior (Lanovaz et al., 2012). To investigate this relationship, the researchers evaluated each participants' preference on five different songs using a musical preference assessment. From this assessment, the researchers chose a highly preferred and lowly preferred song to use in a multielement treatment package. While both types of music lowered the levels of vocal stereotypy, highly preferred music had a larger effect in all but one participant. Furthermore, every participant engaged in higher levels of vocal stereotypy during the no-interaction condition than in the music conditions. One interesting finding Lanovaz recorded was while the vocal stereotypy may have decreased, the on-task behavior did not necessarily

increase. It may be that there are alternative behaviors competing with the participants' on-task duration while the music was playing. Their explanation is that the participants would orient towards the speakers and engage in higher levels of motor stereotypy. The current study will attempt to avoid this outcome through the use of headphones and the implementation of RIRD if matched stimulation by itself is not effective.

Response Interruption and Redirection

Response interruption and redirection is one common treatment to decrease vocal stereotypy (RIRD; Ahearn et al., 2007). RIRD is a type of consequence intervention which includes delivery of demands whenever vocal stereotypy occurs. That is, the individual is redirected to emit three correct, typically vocal, responses without the presence of vocal stereotypy. These researchers were the first group of scientists to use RIRD on vocal stereotypy with a similar population of children on the autism spectrum. They had a familiar teacher state demands that the children had already complied with during academic programming. However, these vocal demands were contingent on vocal stereotypy. All four participants were recorded as having significantly lower levels of stereotypy when compared to baseline. Furthermore, three of the four children had higher levels of appropriate communication (Ahearn et al., 2007).

Researchers have evaluated whether RIRD functions as an extinction procedure by interrupting the automatic reinforcement contingency or a punishment procedure (Ahearn et al., 2011). Ahearn and colleagues found that forms of RIRD that did not interrupt the automatic reinforcement contingency (i.e., RIRD with motor tasks) also resulted in decreased vocal stereotypy for four participants. In addition, RIRD also typically involves brief removal of preferred items in order to deliver demands, which may function as negative punishment (i.e., response cost). Thus, the functionality of RIRD can be argued to be a punishment procedure.

There are many procedural variations of this procedure, and most had similar levels in the deceleration of the targeted behavior. There are also areas of RIRD which do not have a standardized or best practice, such as the optimal instances of demands needed to decrease the vocal stereotypy. Researchers compiled a meta-analysis of RIRD studies, and while most studies had promising results, there were more areas that lacked data in important areas such as social validity measures and treatment integrity measures (Martinez et al., 2013).

RIRD and MS

There are few studies that have examined a treatment package of MS and RIRD. The first of these studies came from Love and colleagues. The MS portion of this package was a preferred auditory toy. The RIRD portion included the removal of this toy contingent upon the onset of vocal stereotypy. Redirection came in the form of saying the child's name and then having the child echo a predetermined word. While all interventions decreased the stereotypy, MS+RIRD had the biggest decelerative effect (Love et al., 2012). However, this study had some flaws and limitations. One of the biggest limitations is that the stimuli used for the matched stimulation were removed contingent on vocal stereotypy, which weakens the connection between MS and RIRD and their combined influence on the duration and rate of the behaviors. This study also had interrupted data collection, meaning while RIRD was removed, data were not collected or were excluded from the RIRD implementation. The study also only took place in the clinical setting with staff, undermining the generality of the intervention.

A follow-up study by Gibbs et al. (2018) replicated and extended the Love and colleagues' study by using an uninterrupted form of data collection. They also tested these interventions during task demands rather than during play times to determine the interventions' effectiveness on increasing on-task behavior. Gibbs noted that continuous access to preferred

auditory toys may have adverse effects on on-task behavior. For this reason, they evaluated non-contingent music instead. Furthermore, they addressed the generality issue by assessing the treatments when given by caregivers in different settings. However, this study still had some limitations that need to be addressed. The major limitation of this study is that there was not an evaluation of matched stimulation by itself. In addition, the researchers did not include a baseline or control condition. A comparison of MS alone, RIRD alone, and the package of the interventions would show if MS through noncontingent music by itself would have an abative effect on vocal stereotypy.

Treatment studies that target vocal stereotypy will typically state that the behavior is being targeted for reduction because it interferes with the participants ability to learn. However, few provide objective measures to determine when it is appropriate to intervene. Researchers first evaluated the extent to which treatment procedures above and beyond typical classroom behavior management were necessary to decrease vocal stereotypy. If additional procedures were needed, they progressed from less to more intrusive interventions (Cook et al., 2018). Following a screener for automatic reinforcement, researchers evaluated the addition of academic tasks such as puzzles or other similar activities. If this did not lower the stereotypy, they moved on to the effects of continuous music. The model then progresses by combining the music with the tasks. The final step in their model has two subsections, one of them being positive practice overcorrection, while the other section is a differential reinforcement of alternative behavior with edibles. Their results showed significant decreases in vocal stereotypy when music was introduced into their progressive model. One limitation of this study is that the researchers did not follow the steps of the model precisely for each participant, with one participant having

multiple phases combined instead of the original component analysis structure they created. The current study will be evaluating the same treatments and the same order for each participant.

The current study seeks to answer the limitation posed by Gibbs (2018) as well as extend its procedures. To address previous study limitations, we included a condition of matched stimulation by itself before progressing to more intrusive interventions, as this approach is considered best practice. We also included the addition of a social validity score to determine which intervention or package is most valid. RIRD itself can be considered a punishment procedure, and we have an obligation to give our clients the least intrusive and aversive treatment. Furthermore, we recorded data on appropriate behavior during the experimental sessions to determine if these interventions will concurrently accelerate on-task behaviors. Procedurally, we used headphones to play music, similar to previous research. We also implemented these interventions during task-demand conditions, using uninterrupted data recording, and assessed the generality of the outcomes.

Chapter 2

Methods

Participants

Participants included in this study were three individuals with ASD who were receiving early intervention applied behavior analysis services from the Scott Center for Autism Treatment, a university-based center for individuals with an ASD diagnosis. All participants were referred to the study by their case managers due to high levels of vocal stereotypy which were negatively affecting treatment progress.

Nemo was a 4-year-old boy whose vocal stereotypy included repetitive scripting of videos, repeating nonsense sounds, and screaming. VS occurred across a variety of activities ranging from academic tasks to free play. Furthermore, VS occurred at a frequency and magnitude that affected daily living skills and time spent completing academic tasks. Nemo communicated vocally, typically in one-to-four-word sentences. Furthermore, Nemo had two-to-four-word mands such as “want to go outside” and “want play doh.”

Kent was a 4-year-old boy whose vocal stereotypy included repetitive scripting of videos, repetitive scripting of therapist and caregiver prompts, and repeating nonsense sounds. Kent also breathed through his mouth loudly; this issue was the subject of doctor’s visit and was not included as a form of vocal stereotypy. Kent had strong vocal communication skills but was rigid with his responses. Variability in responses, such as “I’m good,” “I’m alright,” were targeted in his ABA therapy sessions. VS occurred in a variety of environments ranging from free play to academic tasks. Kent was recruited for participation because his VS occurred most often in academic settings such as a group learning environment provided at his therapy site. His VS

affected his ability to attend to tasks and complete worksheets within a reasonable timeframe. Furthermore, his scripting during group activities distracted his peers.

Leia was a 4-year-old girl whose vocal stereotypy included repetitive scripting of videos and caregivers, repeating nonsense sounds, and screaming. Leia has limited vocal communication, speaking in sentences less than 5 words long in most cases. Similarly to Kent, Leia was chosen for participation due to the negative affects her VS had during group and individual academic activities.

We conducted a screener functional analysis on each participant to ensure vocal stereotypy is maintained by automatic reinforcement. We also measured the levels of vocal stereotypy during academic tasks to ensure vocal stereotypy occurs at high rates in this context. Participants all exhibited the ability to wear and tolerate over-the-ear headphones for at least 5 minutes.

Materials and Settings

Sessions were conducted in individual treatment rooms separate from the clients' typical classroom. The rooms were 10 by 10 feet, and one wall had a one-way window for observation. Within the room there was a table, two chairs, and task materials such as puzzles, sequencing cards, and worksheets. In the matched stimulation (i.e., music) condition, there was a music playing device and a set of headphones. Decibel levels were monitored throughout the experiment, with the volume of the music never exceeding 70 decibels (WHO-ITU, 2019). If severe problem behavior, including aggressive or self-injurious behavior were to occur at any point in the study, the session was stopped de-escalation procedures were implemented, however, this did not occur.

Chapter 3

Data Analysis

Response Measurement and Definitions

All sessions were 5 min in length. In each 5 min session, the percentage of session with vocal stereotypy and on-task behavior was recorded using a computerized data collection program called BDataPro (Bullock et al., 2017). This system allows for continuous data collection and provides summary data for each variable. Session duration data were collected on two dependent variables: vocal stereotypy (i.e., immediate onset and offset) and on-task behavior (3-sec onset and 3-sec offset). Furthermore, for participants who experience the RIRD intervention, the total duration (i.e., onset and offset) of RIRD implementation was recorded.

Vocal stereotypy is the instance of contextually inappropriate vocalization lasting at least 3 seconds (Gibbs et al., 2018). Examples of VS include repetitive sounds that are contextually inappropriate, as well as inappropriate babbling, singing, laughing, delayed echolalia (i.e., scripting previously heard phrases) or noises such as squeals. Negative vocalizations (i.e., whining, crying), appropriate vocalizations, or contextual singing were not recorded as a part of the stereotypy.

Singing is the instance of contextually appropriate vocalization lasting at least 3 seconds. In the case of this study, singing was determined by cross referencing the vocalizations the participant made with the song that was currently playing through the headphones. It is important to note that if the headphones were not on, a different song was playing, or no music was playing, we would count these vocalizations as VS due to its noncontextual nature.

On-task behavior is defined as appropriately sitting and engaging with task materials. On-task behavior was scored when the participant was oriented and engaged in the task given to

them, as well as the manipulation of the task materials with their intended use. For example, on-task behavior during puzzle completion would include sorting puzzle pieces and moving pieces to set location. Simply holding a puzzle piece while oriented away from the other materials would not be scored as on-task. On-task behavior was recorded throughout the session and was scored when the participant was engaged for at least 3 seconds. The duration measure was stopped when the participant was not engaged for 3 seconds.

For individuals who experienced RIRD, the duration of the RIRD implementation was recorded by pausing the session time in BDataPro when the interruption occurs, and restarting the session time when the participant is redirected back to the independent task. The data collection program provides the total pause and session time in the summary output. Note the overall session time remained constant at 5 minutes. Procedural fidelity data were collected on matched stimulation and RIRD implementation. Outside observers reviewed video recordings of at least 33% of the sessions that will be conducted. The observers completed a checklist which included questions on correct use of MS procedures and RIRD implementation.

Interobserver Agreement

A second observer trained on the response definitions recorded data for at least 33% of the sessions in all conditions for each participant. Partial agreement within intervals was calculated by dividing the session into 10-s intervals and assessing average agreement within the interval by dividing the smaller number by the larger number and multiplying by 100 to yield a percentage (e.g., Mudford et al., 2009). For example, if Observer A scored target behavior as occurring for 6s in an interval and Observer B scored target behavior as occurring for 10s seconds in the interval, the agreement for that interval would be 60.0%. Agreement on the nonoccurrence during an interval was calculated as 100% agreement.

IOA was collected for 46.8% of sessions for Nemo, 33.3% of sessions for Kent, and 48.6% of sessions for Leia. For Nemo, IOA for vocal stereotypy was 91.3% (range: 82.1-100.0%) and IOA for on-task behavior was 85.3% (range: 78.3-94.3%). For Kent and Leia, IOA data collection is currently in progress and will be updated shortly.

Extended No-Consequence Screener

An extended enriched no-consequence screener procedure was conducted to determine whether the vocal stereotypy persists without social consequences. Sessions were 5 min in length and at least three sessions were conducted or until stable response patterns were observed. In this screener, both the participant and experimenter were in a treatment room. The participants were provided moderately preferred items which were reported to evoke vocal stereotypy. The experimenter withheld all social interaction, and no programmed consequences were delivered contingent on vocal stereotypy.

Pre-Assessments

The selection of independent tasks (e.g., puzzles, worksheets) for each participant were chosen by interviewing therapists and caregivers. We also asked caregivers to provide a list of at least 5 songs to include on the MS playlist. If MS alone was ineffective at reducing vocal stereotypy, we moved to RIRD. We determined RIRD task demands by interviewing staff members and conducting probes for compliance. Demands included sound imitation, word imitation, intraverbal fill-ins and questions. For individuals who do not have vocal imitation in their repertoire, RIRD demands included motor imitation tasks. Demands with at least 80% compliance on 5-trial probes were included in the RIRD procedure.

Experimental design

The effects of MS on the percentage of session with vocal stereotypy and on-task behavior was evaluated using a reversal design. For participants who did not show improvement from MS alone, we implemented the package intervention and then a component analysis to determine necessary treatment components. However, to date, we observed reductions with MS alone for both participants.

Procedure

Baseline.

All sessions were 5 min in length, beginning when the independent activity was delivered by the therapist. The therapist provided prompts to engage in the activity once every 30s if the participant was off task. Otherwise, no other programmed consequences were provided. Direct reinforcement for task completion was not provided.

MS only.

MS only sessions were conducted using the baseline procedures described above. Prior to this condition, the experimenter used sound-level meter to measure the volume of the music coming from the headphones. The Decibel reading from the headphones was always within the range of 40 to 45 dB. The experimenter began the session by stating “you get to listen to music while you work today” and placed the headphones on the participant before the task began. If the participant attempted to remove the headphones, they were represented up to five times and then sessions were stopped for the day. However, this did not occur during the current experiment. The music was played on a loop for the duration of the session.

MS+RIRD.

Similar to the MS only condition, the experimenter measured the decibel level of the music coming from the headphones. The experimenter will also provide the same statement when placing the headphones onto the participant. Contingent on vocal stereotypy, the experimenter immediately interrupted the task and initiated RIRD using the procedure described by Love and colleagues. The participant had to emit three consecutive correct responses in the absence of VS to end RIRD implementation. Upon completion of RIRD, the experimenter provided neutral praise and directed the participant back to the task. If a correct response did not occur within 5 seconds of the RIRD demand, or if the vocal stereotypy continued, the experimenter stated the correct answer then represented the demand.

Generalization Assessment

The most effective treatment was assessed during therapist lead (i.e., discrete trial, DTT) instruction for Nemo and during small group independent tasks for Kent to evaluate its effectiveness in other academic settings. We conducted baseline sessions during the specific classroom activities and measured percent of session with vocal stereotypy using procedures described above. We also collected data on therapist instruction delivery and participant correct and incorrect responses to therapist instruction. During all sessions, the participants' current teaching procedures, including prompting, error correction, and reinforcement schedule were in place. For Nemo, we evaluated generalization effects using a reversal design. For Kent, we evaluated generalization effects using a multielement design.

Treatment Preference

We conducted a concurrent operants treatment preference assessment to determine the participant's preference for the intervention. During the treatment preference assessment, two identical independent tasks were presented to the participant. One task was paired with the headphones/music device and the other was not. We alternated pairing as well as location (side) of the headphones. The participant was presented with the tasks and instructed to pick one. Once a selection was made, the client received the task or the task and headphones for 30-45s before the next trial began. A total of 10 trials were conducted for Kent. Following 10 trials of undifferentiated responding for Nemo, we modified the assessment by having music playing from the phone speaker for each trial. If music was chosen, the headphones were plugged in, and he received the task with headphones for 30-45 seconds.

Chapter 4

Results

Figure 1 displays the data for Nemo. The top graph shows the results from the extended no consequence condition. In this condition, Nemo had access to preferred reinforcers for five minutes. In this time, the on-site therapist would provide neutral praise when appropriate, and would not deliver consequences contingent upon VS. In this screener, VS occurred over 80% of the sessions. The middle panel displays results from the treatment analysis. The left graph displays percent of session with stereotypy and the right graph displays percent of session with on-task behavior. During baseline, vocal stereotypy was high, occurring at a mean duration of 70% across 8 sessions and on-task duration had a mean duration of 78% across 8 sessions. During MS, VS was significantly lower, with a mean duration of 1.6% across 8 sessions. On-task behaviors remained at a similar mean level as baseline. The bottom panel displays the results from the generalization assessment. The left graph displays percent of session with stereotypy and the right graph displays percent of session with correct compliance. The procedures were extended to an at desk session with DTT tasks consisting of tacting household objects and family members. In baseline, VS had a mean duration of 43% in 6 sessions and the total mean compliance was 68% over 6 sessions. In treatment, VS was once again significantly lower at a mean duration of 1% across 6 sessions while compliance was slightly lower with a mean duration of 60% over 7 sessions. Possible reasons for the decrease in compliance will be addressed further in the discussion session.

Figure 2 displays data for Kent. The top graph shows the results from the extended no consequence condition. In this condition, Kent was in a stimuli-rich room with a therapist. No consequences were delivered if VS occurred during this assessment. In this screener, VS

occurred over 20% of the sessions. The middle panel graphs show the results from the treatment analysis. The left graph displays percent of session with stereotypy and the right graph displays percent of session with on-task behavior. During baseline, vocal stereotypy was high, occurring at a mean duration of 43% across 8 sessions while on-task duration was variable and had a mean duration of 67% across 8 sessions. During MS, VS was significantly lower, with a mean duration of 3% across 6 sessions. On-task duration was significantly higher with a mean duration of 94% across 6 sessions. The bottom panel displays the results from the generalization assessment. The left graph displays percent of session with stereotypy and the right graph displays percent of session with on-task behavior. Generalization was tested by extending the procedures into a small group setting while using 3-minute sessions. In this setting we recorded VS and on-task duration both with and without headphones on. The data show that levels of VS remained near zero levels across the 6 sessions when headphones were used. During baseline, the VS initially started much higher at around 40% of the first trial. However, we observed that the levels decreased over subsequent sessions. Although data were more variable in baseline sessions, there was a marked difference between baseline and treatment conditions.

Figure 3 displays Leia's results. We started by conducting an extended no-consequence screener in which the trials were 5 minutes long and held in an stimuli-enriched environment. During the screener, vocal stereotypy was consistently over 40% of the trials. The middle panel graphs show the results from the treatment analysis. The left graph displays percent of session with stereotypy and the right graph displays percent of session with on-task behavior. During baseline, vocal stereotypy was high, occurring at a mean duration of 31.5% across 9 sessions while on-task duration was variable and had a mean duration of 81.5% across 9 sessions. During MS, VS was significantly lower, with a mean duration of 5.5% across 6 sessions. On-task

duration was marginally higher with a mean duration of 90.2% across 6 sessions. The bottom panel displays the results from the generalization assessment. The left graph displays percent of session with stereotypy and the right graph displays percent of session with on-task behavior. Generalization was tested by extending the procedures into a small group setting while using 3-minute sessions. In this setting we recorded VS and on-task duration both with and without headphones on. The data show that levels of VS remained near zero levels across the 6 sessions when headphones were used. During baseline, the VS initially started much higher at around 90% of the first trial. On-task duration showed initially low levels during music implementation, but steadily increased across subsequent trials whereas baseline levels were variable with no clear trend.

Figure 4 displays the results from the concurrent operants treatment preference assessment for all participants. For Nemo, responding was variable, with no clear indication if music or no music was more preferred. We modified the way the assessment was conducted by having music being played from the speakers during selection process, and if music was chosen the headphones were plugged in for listening. When music was played through the speakers when running the assessment, he chose music 100% of trials. For Kent, music was chosen over no music 100% of trials. When testing Leia, the no-music condition was chosen during the first trial, but every subsequent trial she chose the music condition resulting in a 90% cumulative choice.

Chapter 5

Discussion

Autistic individuals and individuals on the autism spectrum commonly exhibit vocal stereotypy, which refers to nonfunctional or noncontexual vocalizations. This behavior often persists in the absence of social consequences, as it is being maintained by automatic reinforcement (Vollmer et al., 1994). Because of the sensory consequences of the behavior, an individual may engage in high levels of stereotypy, to the detriment of other more socially appropriate and adaptive skills. For instance, Nemo could not wash his hands appropriately because his VS was extremely intrusive. For those like Nemo, this absence in attending due to VS could be potentially dangerous. One example of this is if the stereotypy competes with the attending to environmental cues, such as watching for vehicles when crossing the road.

When vocal stereotypy causes delays in skill acquisition or is disruptive to the environment, clinicians may implement treatments including competing stimuli, differential reinforcement of other behavior, and consequences to decrease the behavior. Although competing stimuli in the form of music has been successfully implemented (e.g., Saylor et al. 2012), previous research has shown that a combination of auditory stimulation and response interruption and redirection, a form of punishment, is more effective (e.g., Gibbs et al., 2018). Our investigation aimed to implement a least to most intrusive intervention by first evaluating matched stimulation and then adding RIRD if necessary. We found that playing music through headphones as matched stimulation significantly decreased vocal stereotypy for all participants across multiple settings. Additionally, there were marginal increases in on-task behavior. To date, RIRD was not a necessary component for effective treatment, showing less intrusive interventions could be implemented and still show promising results.

Researchers have evaluated if providing access to noncontingent stimuli, especially stimuli that may match the purported sensory consequences, is effective at reducing automatically reinforced stereotypy (e.g., Piazza et al., 1996). One of the purported sensory consequences for vocal stereotypy is the sound produced by the behavior. Thus, the effects of matched stimulation with auditory stimuli on vocal stereotypy began with a free-operant comparison between auditory toys and non-auditory toys (Taylor et al., 2005). During this initial study, it was found that the toys that produced an auditory stimulus also subsequently lowered the levels of vocal stereotypy for all participants. Researchers also used auditory toys in their treatment, and they found that the matched stimulation may have had lingering abolishing effects that lowered VS during baseline trials as well (Rapp et al., 2007).

Researchers evaluated noncontingent access to music, a white noise recording, and recordings of the participants' own vocal stereotypy on levels of vocal stereotypy across participants (Saylor et al., 2012). This was one of the first studies to use music by itself as a form of matched stimulation, as well as recordings of the participants' own VS. Both the music and recorded VS lowered the levels of VS in the participants, but the music treatment had higher scores on social validity. Later, researchers used non-contingent access to music with uninterrupted data recording methods (Gibbs et al., 2018). However, this study did not include a baseline condition, nor a control condition. Furthermore, the study did not see if MS by itself would have been effective before moving on to more intrusive interventions. As the current study has shown, MS in the form of noncontingent access to music by itself can be an effective way to both lower stereotypy and increase on-task duration.

The data from Nemo follow the results from similar studies on noncontingent access to auditory stimuli. The extended play pre-assessment showed high levels of vocal stereotypy, as

did the first puzzle baseline. When headphones were introduced, the levels of vocal stereotypy significantly decreased, with some sessions having near-zero levels. The procedure was extended into an at-desk discrete trial training setting to test for generalized results. What was found during these sessions is that without the headphones, vocal stereotypy and non-compliance were at levels similar to the puzzle baseline. When headphones were introduced, the levels of vocal stereotypy once again lowered significantly, but on-task behavior was near baseline levels and was highly variable. Although patterns of on-task behavior were similar, the topography was markedly different.

Anecdotally, in baseline sessions, the non-compliance consisted of high levels of vocal stereotypy, standing up from the desk, trying to grab objects, aggressions, and some instances of negative vocalizations. During the intervention sessions, the aforementioned behaviors were significantly lower and the decrease in on-task behavior was almost entirely due to not responding to the prompts from the therapist. It appeared that non-responding was due to the participant being engaged listening to the song, which may have competing with other reinforcers for compliance.

This intervention closely resembles other appetitive antecedent options that enrich the environment. Environmental enrichment was first described by Horner (1980) to decrease self-injurious behaviors and aggressions. In the seminal study, Horner gave free access to appetitive stimuli independent from any behavior emitted at that time. While this did have the intended effect of lowering the abhorrent behaviors in those individuals, when a consequence was added such as a differential reinforcement component, the efficacy of the intervention increased significantly (Horner et al., 1980). We believe that if a consequence component was implemented for Nemo, we may have seen an increase of compliance during generalization.

From a procedural standpoint, the only difference that would have to change is that the headphones would either be turned off or taken away while the therapist prompted a learning objective. For Nemo, he would receive the headphones contingent upon a correct response, but for learners whose main pitfalls in learning may include not responding at all, headphones could be given back for any response.

In a meta-analysis of environmental enrichment studies, only 13% of all the studies were conducted on vocal stereotypy (Gover et al., 2019). On top of this low percentage, there was only one mention of music as an intervention, and the only example given increased the automatically reinforced behavior rather than lowering it. However, in that case the behavior was object twirling, which does not match the type of auditory stimulation that music provides. The current study shows that music can be used as a form of environmental enrichment and is effective in both lowering the targeted behavior and increasing alternative behaviors. Given the results we found with this form of enrichment without specifically programming for it, future research should contrive the environment to better test the efficacy of music as environmental enrichment.

Another area of interest is changes to the occurrence and topography of compliance. When compliance with therapist instructions did occur, it appeared to have a longer latency to responding, and some targets had a stutter. We theorize that this may be due to the words of the song itself competing with engaging in the target response. However, additional sessions with more sensitive data collection procedures may be necessary to determine underlying causes of these differences. From a validity standpoint, we can look back through the video recordings of the sessions to determine if there were any significant changes in affects and other secondary behaviors. As noted above the biggest difference in these behaviors came from Nemo. We believe that music had this effect on Nemo due to the appetitive nature of the intervention.

Kent's data show lower levels of VS in the extended no consequence baseline. However, during the tracing task baseline, the VS was much higher. These data support caregiver report of VS interfering with academic tasks and indicate that an intervention is needed. When headphones were introduced, VS decreased to near-zero levels instantly. Operational definitions for singing were created as this participant sang along to portions of the songs that were playing. Although singing was observed, it still occurred at lower levels than VS with a max of 13% of sessions compared to 40% for VS. Singing along to a song as a variable was not seen in other studies involving vocal stereotypy, yet all three participants sang as a part of their vocal stereotypy.

We observed different response patterns for Kent in the generalization assessment. During the small group tracing task, we alternated baseline and MS sessions. The data in baseline sessions were lower overall and demonstrate a decreasing trend in VS after the first session with music. It is hypothesized that the shorter, more rapid switching between conditions may have led to carryover effects. Previous research has shown similar results when using auditory stimuli as a form of matched stimulation, lending to the hypothesis that music may have an abolishing effect on VS even after the stimuli is removed (Rapp et al., 2007).

During implementation for Leia, we saw similar results of near-zero levels of VS in both the original treatment and generalization sessions. For Leia, we observed that even while music was playing, she engaged in noncontextual singing (i.e., she was singing a completely different song than the one that was playing). However, singing in general still occurred for longer durations when the headphones were not present.

Clinical Implications

The current data highlight the need to evaluate less intrusive interventions prior to implementing consequence-based procedures. Matched stimulation was effective for both participants in the absence of more intrusive procedures and may have additional benefits. For example, using noncontingent access to music requires less response effort to implement than RIRD. Furthermore this intervention is antecedent based rather than consequence based, leading to proactive steps taken rather than reactive. Additionally, MS operates by adding purportedly preferred stimulation to the environment rather than delivering aversive consequences (i.e., demands) contingent upon stereotypy. The procedural fidelity of MS is also much simpler than that of the RIRD procedure. Noncontingent access to music as a form of matched stimulation may be more socially valid to implement outside of the clinic as well.

When comparing music to other forms of auditory stimulation such as white noise and self-recordings of stereotypy, music scored the highest on social validity surveys (Saylor et al., 2012). While we did not compare music to those forms of stimulation in the current study, we did compare preferred music to non-preferred music with Leia. We had started the music intervention with little information about what types of music she preferred, so we started with instrumental versions of songs from popular children's movies. What we found is that when the music was novel or not preferred, Leia's vocal stereotypy was higher, and her attending was lower. When preferred music for her was found, we found that VS levels decreased to near zero levels. Future research in this area should look closer into difference in VS levels when comparing novel music to preferred music.

An area of research that has been opened through this study is singing as a behavior. As mentioned above, Kent sang along with small portions of the songs played in through the

headphones. He was the only participant in our study to do so and the only participant within the auditory stimulation studies that preceded this one. Singing is a form of music, and in some cases, it is more socially valid than instrumental music. Yet in autistic children, we rarely see contextual singing, rather we see it as a part of their vocal stereotypy scripts. There are multiple areas of singing that can be examined, starting with its function. Discovering why we observe singing within vocal stereotypy but not in a contextual setting would also be an important area of research.

Limitations

A limitation of the study is that we did not investigate the longevity of the effectiveness of this intervention. Future research should investigate the long-term effects of matched stimulation. It is possible that satiation to the MS intervention could occur over time. Furthermore, future research should investigate ways to mitigate satiation effects, such as the use of different music types. A major limitation that was brought forth through the completion of this intervention is that wearing headphones is not feasible for some individuals within this population. However, headphones were chosen for this study because they have the ability to present auditory stimulation without disrupting others in the environment. For individuals who cannot or will not wear headphones, MS may be delivered in other ways such as a speaker at low volume near the individual's work area. Future research should compare the effectiveness of methods of implementing MS on levels of vocal stereotypy.

Due to time constraints, social validity scores were unable to be recorded for caregivers of the participants. While it is unlikely that this intervention is not social valid due to the literatures' previous findings such as Love (2012) and Gibbs (2018), it is still a limitation of this study. Ideally, caregivers would have sat for an observation to view how the intervention was

conducted and given an opportunity to implement the procedures. After this, caregivers would have been given a survey modified from Love (2012), and results would have been scored on a Likert scale of 1 to 5.

On two occasions during intervention, the headphones slipped off the participant's head and lead to an outlier of on-task duration in both Leia and Nemo. These pitfalls show the dynamic natural of wearable items, especially with individuals who move around a lot or fidget. We also had technological issues, including music cutting out on one occasion. Future implications of music as an intervention should investigate a specialized music platform that is dependable and easier to control. Furthermore, the headphones we used were connected to a phone via a wired connection. While this did not pose a major distraction for our participants, on some occasions they did fidget with it. Future implications of this intervention may find it better to use Bluetooth options, as there is not a chord that could provide stimulation to the participant listening to music.

Headphones also may have the added benefit of blocking out other noises. For example, during generalization sessions for Nemo, he did not orientate towards the loud noises in the room that he usually would be distracted by. This intervention may be well suited and have additional benefits for individuals who are "overstimulated" by noises in typical environments. We evaluated levels of vocal stereotypy and on-task behavior. It may be possible to collect data on other dependent variables (e.g., participant affect, startle responses, ear covering) to assess other potential advantages.

Lanovaz and colleagues noted that while vocal stereotypy decreased in their study, on-task duration did not show major increases or decreases. While we found differing results in two of our participants, Nemo closely followed the results found in their study (Lanovaz et al., 2012).

As discussed earlier, we found that the topography of behaviors between the music and no music trials were markedly different, and that compliance did not increase or decrease. However, what the data does not show is the total time it took to gain his attending and for how long we maintained his attention. Without headphones, gaining his attention was nearly impossible and when attending was gained, it was fleeting. When headphones were introduced, attending and eye contact improved, but the therapist had to repeat prompts before a response was given. Procedural changes to the intervention such as making the music play contingent upon a response should be investigated to determine if on-task performance could increase further.

References

Ahearn, W. H., Clark, K. M., MacDonald, R. P., & Chung, B. I. (2007). Assessing and Treating Vocal Stereotypy in Children with Autism. *Journal of Applied Behavior Analysis. 40* (2), 263–275. https://doi.org/10.1901/jaba.2007.30-06

Ault, M. J., & Griffen, A. K. (2013). Teaching with the System of Least Prompts: An Easy Method for Monitoring Progress. *Teaching Exceptional Children*, *45*(3), 46–53. https://eric.ed.gov/?id=EJ1008512

Bullock, C. E., Fisher, W. W., & Hagopian, L. P. (2017). Description and Validation of a Computerized Behavioral Data Program: "BDataPro." *The Behavior Analyst*, *40*(1), 275–285. https://doi.org/10.1007/s40614-016-0079-0

Cook, J. L., & Rapp, J. T. (2018). To What Extent Do Practitioners Need to Treat Stereotypy During Academic Tasks? *Behavior Modification*, *44*(2), 228–264. https://doi.org/10.1177/0145445518808226

Gibbs, A. R., Tullis, C. A., Thomas, R., & Elkins, B. (2018). The effects of noncontingent music and response interruption and redirection on vocal stereotypy. *Journal of Applied Behavior Analysis*, *51*(4). https://doi.org/10.1002/jaba.485

Gover, H. C., Fahmie, T. A., & McKeown, C. A. (2018). A Review ff Environmental Enrichment as Treatment for Problem Behavior Maintained by Automatic Reinforcement. *Journal of Applied Behavior Analysis*, *52*(1), 299–314. https://doi.org/10.1002/jaba.508

Hagopian, L. P., Rooker, G. W., & Zarcone, J. R. (2015). Delineating subtypes of self-injurious behavior maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*, *48*(3), 523–543. https://doi.org/10.1002/jaba.236

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*(2), 197–209. https://doi.org/10.1901/jaba.1994.27-197

Lanovaz, M. J., Rapp, J. T., & Ferguson, S. (2012). The Utility of Assessing Musical Preference before Implementation of Noncontingent Music to Reduce Vocal Stereotypy. *Journal of Applied Behavior Analysis*, *45*(4), 845–851. https://doi.org/10.1901/jaba.2012.45-845

Lovaas, I., Newsom, C., & Hickman, C. (1987). Self-stimulatory behavior and perceptual reinforcement. *Journal of Applied Behavior Analysis*, *20*(1), 45–68. https://doi.org/10.1901/jaba.1987.20-45

Love, J. L., Miguel, C. F., Fernand, J. K., & LaBrie, J. K. (2012). The effects of matched stimulation and response interruption and redirection on vocal stereotypy. *Journal of Applied Behavior Analysis*, *45*, 549-564. https://doi.org/10.1901/jaba.2012.45-549

Martinez, C. K., & Betz, A. M. (2013). Response Interruption and Redirection: Current Research Trends and Clinical Application. *Journal of Applied Behavior Analysis*, *46*(2), 549–554. https://doi.org/10.1002/jaba.38

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A Brief Functional Analysis Of Aggressive and Alternative Behavior in an Outclinic Setting. *Journal of Applied Behavior Analysis*, *24*(3), 509–522. https://doi.org/10.1901/jaba.1991.24-509

Piazza, C. C., Adelinis, J. D., Hanley, G. P., Goh, H.-L., & Delia, M. D. (2000). An Evaluation ff the Effects of Matched Stimuli on Behaviors Maintained by Automatic Reinforcement. *Journal of Applied Behavior Analysis*, *33*(1), 13–27. https://doi.org/10.1901/jaba.2000.33-13

Piazza, C. C., Fisher, W. W., Hanley, G. P., Hilker, K., & Derby, K. M. (1996). A Preliminary Procedure for Predicting the Positive and Negative Effects of Reinforcement-Based Procedures. *Journal of Applied Behavior Analysis*, *29*(2), 137–152. https://doi.org/10.1901/jaba.1996.29-137

Piazza, C. C., Fisher, W. W., Hanley, G. P., Leblanc, L. A., Worsdell, A. S., Lindauer, S. E., & Keeney, K. M. (1998). Treatment of Pica Through Multiple Analyses of Its Reinforcing Functions. *Journal of Applied Behavior Analysis*, *31*(2), 165–189. https://doi.org/10.1901/jaba.1998.31-165

Querim, A. C., Iwata, B. A., Roscoe, E. M., Schlichenmeyer, K. J., Ortega, J. V., & Hurl, K. E. (2013). Functional Analysis Screening for Problem Behavior Maintained by Automatic Reinforcement. *Journal of Applied Behavior Analysis*, *46*(1), 47–60. https://doi.org/10.1002/jaba.26

Rapp, J. T., Miltenberger, R. G., Galensky, T. L., Ellingson, S. A., & Long, E. S. (1999). A Functional Analysis of Hair Pulling. *Journal of Applied Behavior Analysis*, *32*(3), 329–337. https://doi.org/10.1901/jaba.1999.32-329

Rapp J. T. (2007). Further evaluation of methods to identify matched stimulation. *Journal of Applied Behavior Analysis, 40*(1), 73–88. https://doi.org/10.1901/jaba.2007.142-05

Safe listening devices and systems: a WHO-ITU standard. (n.d.). Www.who.int. https://www.who.int/publications/i/item/9789241515276

Saylor, S., Sidener, T. M., Reeve, S. A., Fetherston, A., & Progar, P. R. (2012). Effects Of Three Types Of Noncontingent Auditory Stimulation On Vocal Stereotypy In Children With Autism. *Journal of Applied Behavior Analysis*, *45*(1), 185–190. https://doi.org/10.1901/jaba.2012.45-185

Skinner, B. F. (1963). Operant Behaviour. *American Psychologist*, *18*(8), 503–515. https://doi.org/10.1037/h0045185

Taylor, B. A., Hoch, H., & Weissman, M. (2005). The Analysis and Treatment of Vocal Stereotypy In A Child with Autism. *Behavioral Interventions*, *20*(4), 239–253. https://doi.org/10.1002/bin.200

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

Wunderlich, K. L., Hemstreet, R., & Best, L. (2022). A retrospective analysis of stereotypy: Applicability of the Behavioral Subtyping Model. *Journal of Applied Behavior Analysis*, *55*(2), 529–546. https://doi.org/10.1002/jaba.902

Figures

Figure 1

Vocal stereotypy, On-task duration, and compliance for Nemo

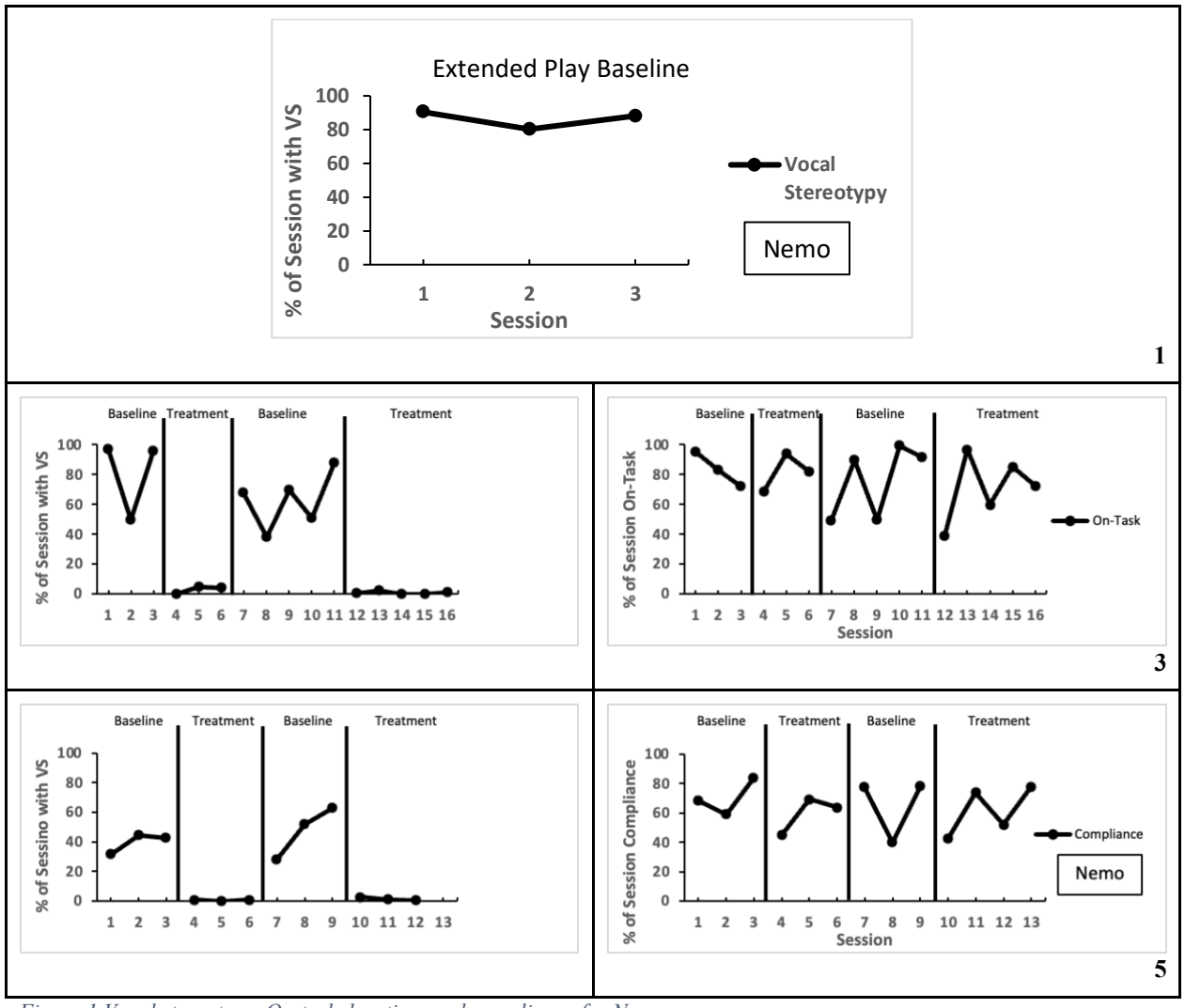


Figure 1 Vocal stereotypy, On-task duration, and compliance for Nemo

Note: Panel 1 Percentage of 10-s intervals Nemo engaged in vocal stereotypy during extended play screener.

Panel 2&3 Percentage of 10-s intervals Nemo engaged in vocal stereotypy and on-task behaviors across sessions.

Panel 4&5 Percentage of 10-s intervals Nemo engaged in vocal stereotypy and compliance across sessions.

Figure 2

Vocal Stereotypy and On-task duration for Kent

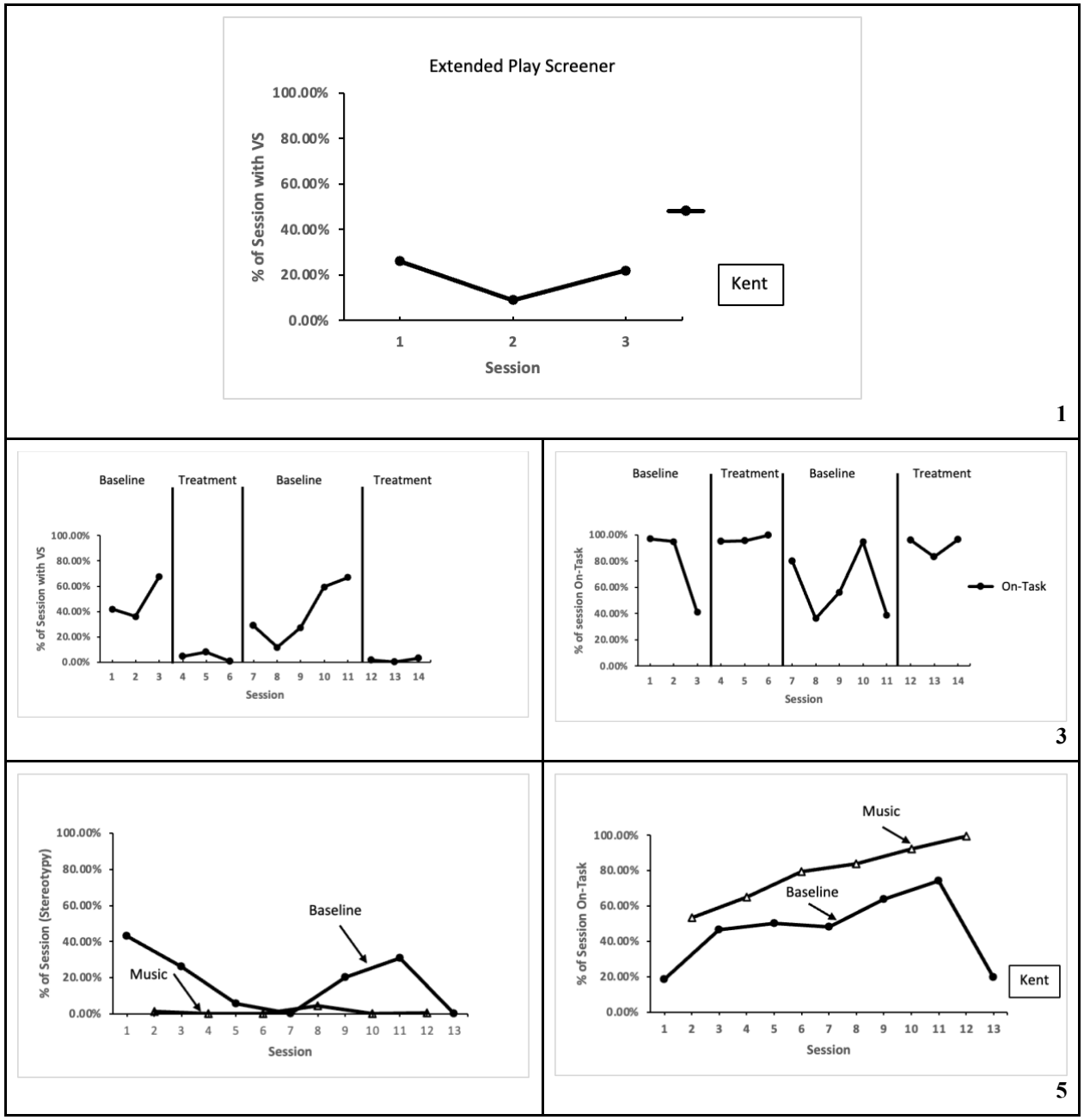


Figure 2 Vocal Stereotypy and On-task duration for Kent

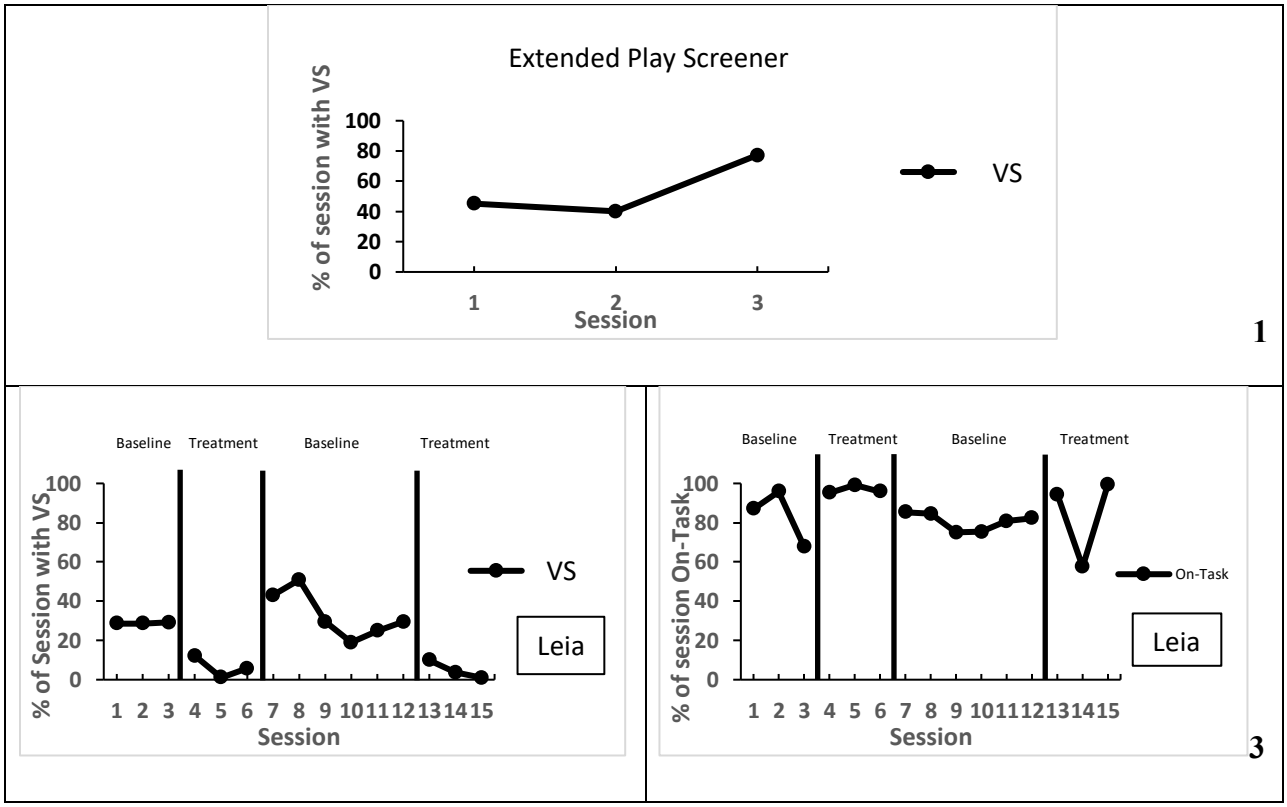
Note. Panel 1 Percentage of 10-s intervals Kent engaged in vocal stereotypy during extended play screener.

Panel 2&3 Percentage of 10-s intervals Kent engaged in vocal stereotypy and on-task behaviors across sessions.

Panel 4&5 Percentage of 10-s intervals Kent engaged in vocal stereotypy and compliance across sessions.

Figure 3

Vocal Stereotypy and On-task duration for Leia



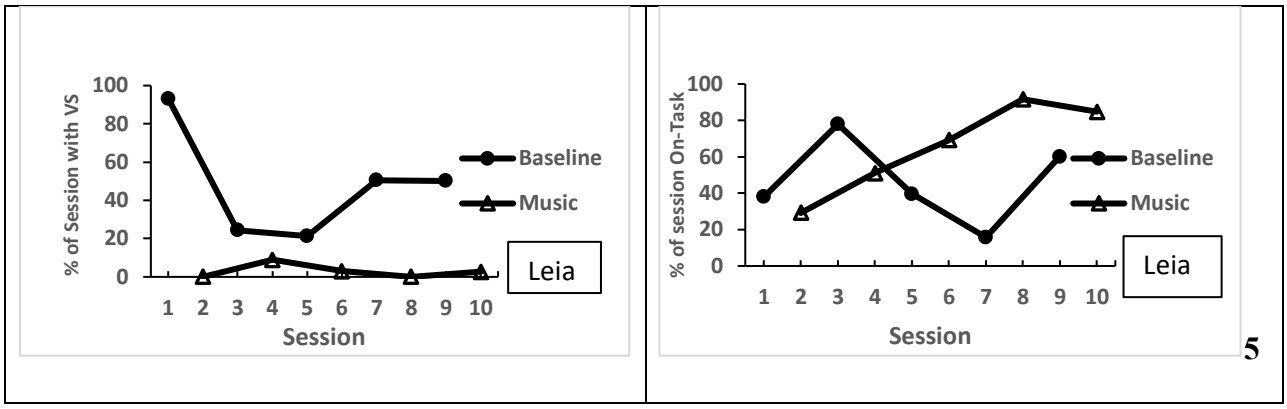
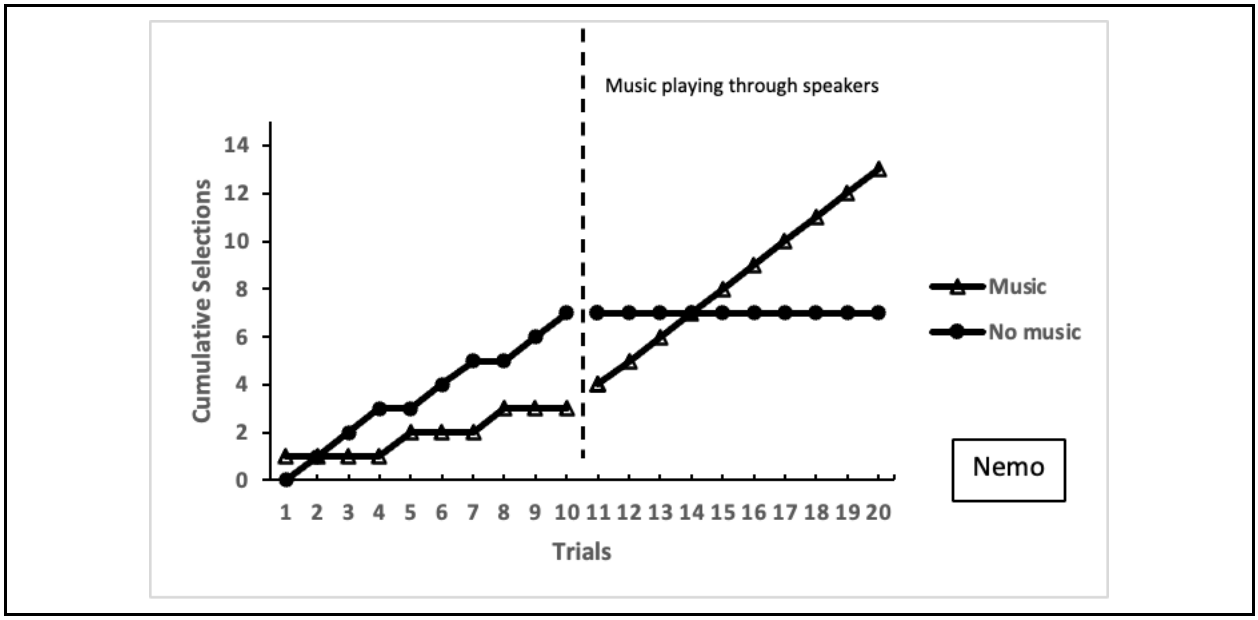


Figure 3 Vocal Stereotypy and On-task duration for Leia

Figure 4

Treatment Preference results



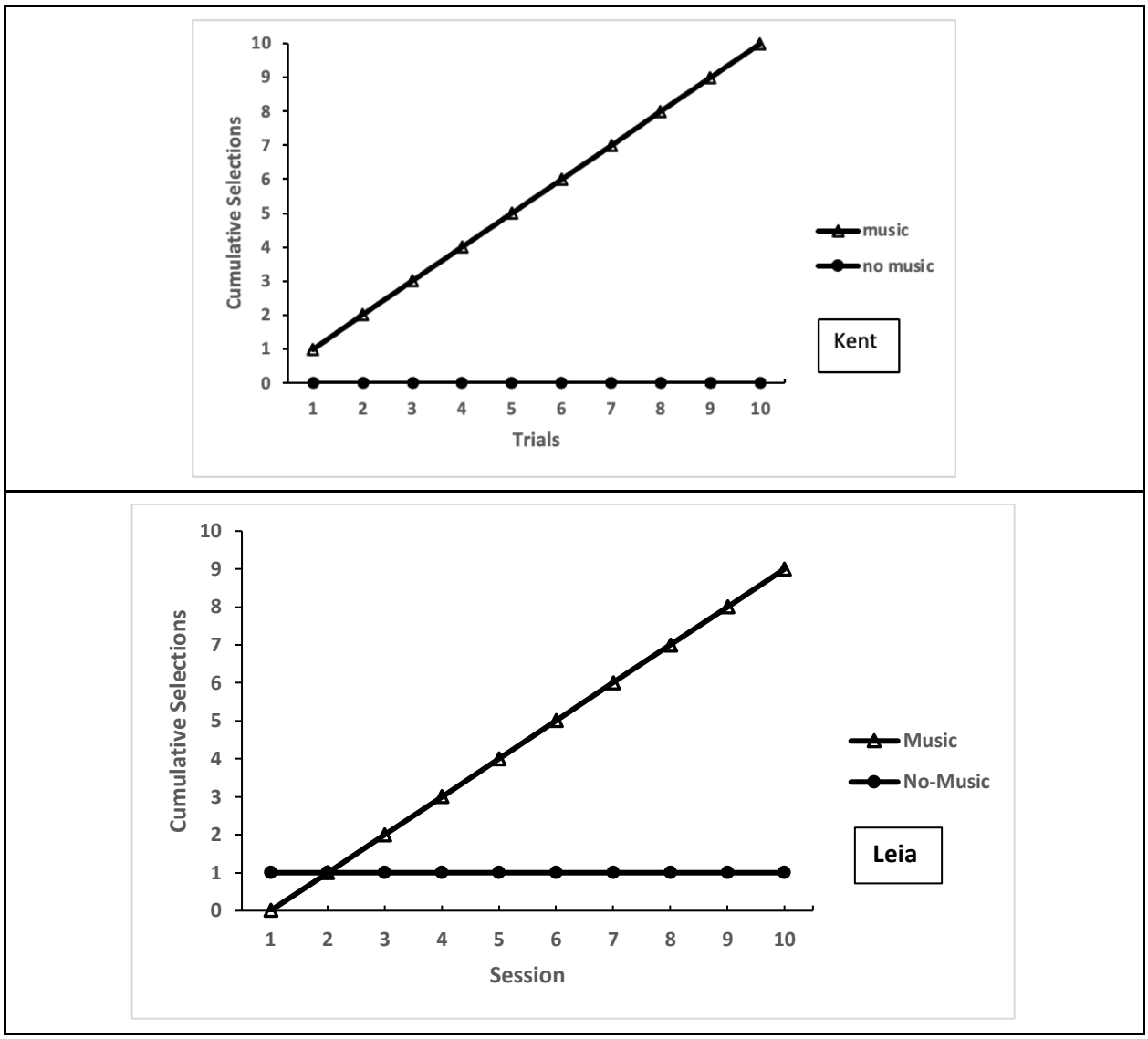


Figure 4 Treatment Preference results

Note: Cumulative selections between music and no music from participants