Subtyping Stereotypy and Delineating Effective Type-Based Treatment

by

Katherine Therese Haggerty

A thesis submitted to the School of Behavior Analysis of
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master's of Science
in
Applied Behavior Analysis &
Organizational Behavior Management

Melbourne, Florida
July, 2019

We the undersigned committee hereby approve the attached thesis, “Subtyping Stereotypy and Delineating Effective Type-Based Treatment” by

Katherine Therese Haggerty

Michael Kelley, PhD., BCBA-D
Professor
School of Behavior Analysis
Thesis Advisor

Sherry Jensen, Ph.D.
Assistant Professor of Economics
Bisk College of Business

David Wilder, Ph.D., BCBA-D
Professor
School of Behavior Analysis

Lisa Steelman, Ph.D.
Professor and Dean
College of Psychology and Liberal Arts

Abstract

Subtyping Stereotypy and Delineating Effective Type-Based Treatment

Katherine Therese Haggerty

Major Advisor Michael Kelley, Ph.D., BCBA-D

Individuals diagnosed with autism spectrum disorder (ASD) often engage in automatically reinforced stereotypic behaviors. These behaviors are often repetitive in nature, lack variability, and persist over time with no identifiable social function. This research used an adapted subtyping procedure to categorize stereotypy based on behavioral patterns in a pairwise functional analysis. Following subtyping, experimenters exposed participants to up to two treatment procedures to determine whether the subtyping procedure is predictive of the efficacy of different treatments. Results from this research showed that stereotypy can be categorized into different subtypes and that those with behaviors categorized into Subtype 1 were responsive to reinforcement based treatment. This information may be used to determine treatment procedures for individuals diagnosed with ASD who engage in stereotypy.

Keywords: stereotypy, subtype, treatment efficacy

Table of Contents

List of Figures

Acknowledgement

I would like to thank my thesis advisor Dr. Michael Kelley for taking me on as a thesis student, and for shaping me into the writer and researcher I am today. I would also like to thank my thesis committee Dr. Sherry Jensen and Dr. David Wilder for both your time and your feedback on my thesis project.

Additionally, I would like to thank each individual who helped me conduct my thesis project. The following individuals helped either run session, collect primary data or IOA, or gave me additional support when I needed it: Justine Henry, Tiara Putri, Jacqueline Noto, Yaara Shaham, Luiz Freitas, Kacie McGarry, Morgan McCoy, Michael Cusick, and John Pruett.

Finally. I would like to thank the Scott Center for Autism Treatment for allowing me to conduct my thesis in the clinic and for supporting me through this process.

Dedication

I would like to dedicate this research project to the individuals and families I have had the privilege of working with, and to those who I will work with in the future. Additionally, I could not be where I am today without the support and guidance from my parents, my siblings, and my grandparents.

...

Subtyping Stereotypy and Delineating Effective Type-Based Treatments

Problem Behavior in Children Diagnosed with Autism Spectrum Disorder

A major concern for those with developmental disabilities is their engagement in serious problem behavior (Rojahn, Matson, Lott, Esbensen, & Smalls, 2001). Children diagnosed with Autism Spectrum Disorder (ASD) often display problem behaviors that can be more dangerous and disruptive than the problem behavior displayed by their typically developing peers (American Psychiatric Association, 2013; Rojahn et al.). For example, individuals with ASD may engage in self-mutilation, often referred to as self-injurious behavior (SIB), which is defined as behavior that produces physical injury to one's own body (Hagopian, Rooker, & Zarcone, 2015; Tate & Barnoff 1996). These individuals also engage in aggression towards others, violent outbursts, and stereotypy (Rojahn et al.). The primary problem when individuals engage in behaviors such as SIB, aggression, and violent outbursts is the potential for physical harm to themselves and others (Rojahn et al.).

Individuals diagnosed with ASD also often engage in other less dangerous, but equally problematic topographies, such as stereotypy. Stereotypy is defined as repetitive behavior, typically motor or vocal, with no social purpose. It lacks variability, is persistent over time, and is immutable when faced with environmental changes (Hagopian et al., 2015; Rapp & Vollmer, 2005; Toper-Korkmaz, Lerman, & Tsami, 2018). Stereotypic behavior often interferes with age-appropriate skill acquisition, and may lead to social rejection and

stigmatization from typically developing peers (Toper-Korkmaz et al.). While these behaviors may not be outright dangerous, they may restrict learning, be stigmatizing, and limit those with developmental disabilities' opportunities to engage with their community (MacDonald et al., 2007, Rojahn et al., 2001).

Maintaining Variables

The majority of human behavior persists due to environmental consequences that strengthen the occurrence of specific behavior, referred to as maintaining variables (Catania, 2013; Cooper, Heron, & Heward, 2007; Skinner, 1953). Cooper et al. defined reinforcement as a stimulus that follows a response that increases the probability of similar responses under similar circumstances. Thus, it follows that both adaptive and maladaptive behavior that continue to occur must be reinforced by some environmental variable. Clinicians and researchers strive to discover the conditions under which responding occurs and does not occur, with the ultimate goal of determining the function of a behavior (including antecedents and consequences that increase the likelihood of a behavior occurring). Voluminous research demonstrates the value of using function-based interventions for effectively treating problem behavior (Beavers, Iwata, & Lerman, 2013; Hanley, Iwata, & McCord, 2003)

Towards that end, clinicians and researchers use functional assessments to better identify the variables maintaining a specific problem behavior. These assessments include both direct and indirect assessments that aim to identify the variables maintaining that problem behavior prior to treatment (Lerman et al.,

1993; Mace & Lalli 1991; Thompson & Iwata, 2007). Indirect assessments use interviews and ratings scales to gather information (Campbell, 2003). Direct assessments include observation of antecedents and consequences to better evaluate what environmental stimuli is maintaining the behavior (Hagopian et al., 2015). Finally, functional analysis of behavior includes formal experimental testing of hypotheses in the context of test-control manipulations of potentially relevant antecedent and consequence variables (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). Treatment for problem behavior based on a functional assessment yields greater decreases in that problem behavior among those diagnosed with ASD (Campbell).

Evolution and Adaptations of Functional Analysis

Iwata et al. (1982/1994) conducted an assessment designed to systematically determine the maintaining variable(s) of a specific behavior, referred to as the function of the behavior. The study included a procedure for assessing the extent to which self-injurious behavior in individuals diagnosed with developmental disorders was a function of one of three variables. These variables were tested using a series of conditions: social disapproval, alone, and academic demand. The *social disapproval* condition tested for positive reinforcement in the form of attention. The *alone* condition tested for self-stimulatory or automatically reinforced behaviors. The *academic demand* condition tested for negative reinforcement in the form of escape. The level of behavior in each condition was compared to a control condition to determine which variable was the most likely

reinforcer for that behavior. This original functional analysis configuration is commonly referred to as the standard FA.

There have been several adaptations to the standard FA since 1982. These include design and measurement modifications, using discriminative stimuli to signal the condition in effect, making adaptations to the session duration or reinforcer duration, and changing the number or order of conditions, among others (e.g., Conners et al., 2000; Fisher, Piazza, & Chiang, 1996; Hammond, Iwata, Rooker, Fritz, & Bloom, 2013; Northup et al., 1991; Thomason-Sassi, Iwata, Neidert, & Roscoe, 2011; Vollmer, Iwata, Duncan, & Lerman 1993). While FAs are typically conducted in an alternating treatment design (also referred to as multi-element), these designs are susceptible to carryover effects, and may compromise interpretation of the results. That is, rapidly alternating conditions could produce undifferentiated responding as the behavior may not readily come under the control of the antecedent and consequence features of the conditions. Vollmer et al. used a reversal design to test whether alternative arrangements might better differentiate the levels of behavior across conditions. Vollmer et al. exposed 4 participants to both multielement FAs and reversal FAs. Results showed that when a multielement FA yields undifferentiated outcomes, a reversal design FA may produce more differentiated responding due to more discriminable conditions. Northup et al. (1991) conducted a brief FA to determine the effects of a truncated assessment procedure when prolonged assessments were impossible. Northup et al. assessed the extent to which a single exposure to each

test condition produced adequate information to implement a treatment. Results suggested that a single exposure to test and control conditions, along with a treatment probe and reversal, was adequate for identifying a function for problem behavior, and a function-based treatment for a small number of individuals.

Other methods to more clearly differentiate FA results include using a fixed order of conditions, including discriminative stimuli in the conditions, and holding reinforcement durations constant across conditions (Conners et al., 2000; Fisher et al., 1996; Hammond et al., 2003). Hammond et al. evaluated the extent to which the order of conditions affected FA efficiency by comparing a fixed condition sequence with a random-sequence FA for 7 individuals. They found that results of the fixed condition sequence yielded outcomes that were more clearly differentiated than when the sequence was randomized. Conners et al. evaluated the addition of discriminative stimuli (SDs) such as a specific therapist and room that indicated the current condition of the FA. They compared FAs using SDs to FAs without SDs and found that there was higher discrimination between conditions when the SDs were included. Fisher et al. evaluated the effects of reinforcer duration on the rate of a target behavior. They found that the duration of access to a reinforcer altered interpretation of FA results. Further, when the duration of reinforcers was unequal across conditions (3 s in attention and play and 30 s in demand and tangible) the results appeared more differentiated than when they are of equal duration (30 s in all conditions).

In addition to these advancements, different measurement procedures or different conditions may be needed to better conduct FAs on specific behaviors or to determine some specific function (Querim et al., 2013; Thomason-Sassi et al., 2011). Thomason-Sassi et al. used *latency to first response* to as an alternative method of measurement in an FA to determine the function of behaviors that cannot occur more than once a session (e.g. elopement or disrobement). They terminated sessions upon the first occurrence of problem behavior such as elopement or disrobement. Thomason-Sassi et al. found correspondence between results of this type of FA and a standard FA in 33 of 38 cases. Another example of adjusting FA procedures to test for a specific behavior or a specific function includes the screening procedures developed by Querim et al. They adapted the functional analysis to specifically test for behaviors maintained by automatic reinforcement prior to the use of multielement FAs to make comparisons between multiple test conditions. The authors used a series of alone or no-interaction conditions as a screening procedure to determine whether behavior was maintained by automatic reinforcement, and then compared those results to a full-fledged multielement FA. They posited that if the behavior is maintained during the initial assessment, then other conditions would not be necessary for determining the function of behavior and implementing treatment. Their results suggested that this screening procedure that took 12.6% of the time of a standard FA was accurate when determining if a behavior was automatically maintained in 93% of data sets. One limitation to this screening procedure is that it determines

whether a behavior is maintained by automatic reinforcement or social variables; but additional conditions are required to determine the maintaining variables if the response is not automatically maintained. In summary, this procedure is recommended for behaviors likely to show automatic reinforcement because this procedure takes less time than a full FA and yields information showing whether the behavior is automatically or socially maintained.

The aforementioned assessment developments are important because FAs are widely used to determine treatment procedures for presenting problem behavior and standard procedures might represent significant barriers to assessment and treatment. Results of voluminous research (see Beavers et al., 2013 and Hanley et al., 2003) suggest that *functional analysis* generally represents a wide variety of procedural variations designed to experimentally isolate operating, maintaining variables and, ultimately, to guide treatment development.

Topographies Likely Maintained by Automatic Reinforcement

Beavers et al. (2013) reviewed 158 studies on the functional characteristics of problem behavior and found that some topographies were more likely to be maintained by automatic reinforcement than others. Responses maintained by automatic reinforcement are different than those maintained by social variables in that the behavior produces its own reinforcement (Hagopian et al., 2015, Toper-Korkmaz et al., 2018). Beavers et al. found that self-injurious behavior was maintained by automatic reinforcement in 29% of cases, while stereotypic behavior was maintained by automatic reinforcement in 82% of cases.

Automatically maintained behavior poses a dilemma for behavior analysts because the reinforcer is not specified and cannot be directly manipulated by a clinician (Piazza, Adelinis, Hanley, Goh, & Delia, 2000; LeBlanc, Patel, & Carr, 2000; Vollmer 1994). In addition to these issues, the reinforcement for automatically maintained behaviors is constantly available to the individual (Piazza et al.).

Prevalence of Stereotypic Behaviors in Children Diagnosed with and without ASD

Stereotypy is defined as behavior that lacks variability, persists over time, and is immutable when faced with environmental changes (Rapp & Vollmer 2005). While stereotypic behavior is a differential diagnostic criterion for ASD, this behavior has also been observed in typically developing infants and children (American Psychiatric Association, 2013; MacDonald, 2007). MacDonald et al. compared levels of stereotypic behavior in children with ASD to levels of stereotypic behavior in their same-age typically developing peers. MacDonald et al. observed that typically developing children aged 2 to 4 engaged in stereotypy between 0% and 27% of the observation period, while children diagnosed with ASD spent between 5% and 61% of the observation period engaging in stereotypic behaviors. They found that while both groups of children engage in this behavior, it becomes less frequent in typically developing children as they mature, while as children with ASD age this time allocation increases (MacDonald et al.).

MacDonald et al. (2007) found that the level of stereotypic behavior in children with ASD clearly differentiates from that of their same age peers as early as age two. Stereotypic behavior often directly interferes with learning either because it makes up a large percentage of individual's behavioral repertoire or is incompatible with behaviors required for attending (children with ASD engage in this behavior 4-8 times more than their typically developing peers; Rapp & Vollmer, 2005; MacDonald). Stereotypy not only affects learning and skill acquisition, but also greatly decreases the likelihood of positive social interactions and can be stigmatizing (Dunlap, Dyer, Koegel 1983; MacDonald).

Treatments of Stereotypy

As stated above, results of functional analyses of stereotypy suggest that this particular topography is often maintained by automatic reinforcement (Beavers et al., 2013). Although the contingency between the behavior and the reinforcer cannot be directly manipulated and controlled, there are several treatments that have been used to successfully treat stereotypic behavior. In a review of behavioral assessments and treatments commonly used to treat stereotypy, Rapp & Vollmer (2005) found that interventions typically rely on procedures that are not based on the operant function and are punishing in nature. These procedures include antecedent manipulations, sensory extinction, differential reinforcement, punishment, and response interruption and redirection (RIRD; Ahearn, Clark, MacDonald, & Chung, 2007; Rapp & Vollmer).

Antecedent manipulations. One type of antecedent manipulation used to reduce stereotypy to clinically acceptable levels is exercise. Kern, Koegel, and Dunlap (1984) exposed 3 participants to mild and vigorous exercise to determine whether these different interventions would decrease levels of stereotypy. Results showed that mild exercise (i.e., playing with a ball) had little to no effect on the level of subsequent stereotypy while vigorous exercise (i.e., jogging), was followed by periods of time with relatively lower levels of stereotypic behavior (Kern et al.). Celiberti, Bobo, Kelly, Harris, and Handleman (1997) replicated and expanded Kern et al.'s results. They exposed a five-year-old diagnosed with ASD who engaged in high rates of automatically reinforced stereotypy to two levels of exercise and measured subsequent levels of two topographies of self-stimulatory behavior (Celiberti et al.). They found that the levels of stereotypy following the high exercise conditions (i.e., jogging) were lower than those following the low exercise condition (i.e., walking) for motor stereotypy, but they were relatively equal for visual stereotypy (Celiberti et al.). Both studies show that children diagnosed with ASD allocated less time to stereotypic behaviors, specifically motor stereotypy, in periods of time immediately following vigorous exercise than in periods of time following mild exercise (Celiberti et al.; Kern et al.).

Environmental enrichment has also been shown to effectively reduce levels of automatically reinforced stereotypic behavior (Berkson & Mason, 1965; Horner, 1980; Piazza, Adelinis, Hanley, Goh, & Delia, 2000). Horner studied the effects of an enriched environment on the stereotypic behavior of 5 residents of an

institution for the mentally disabled. Results from Horner's research showed that levels of stereotypic behavior were reduced in the enriched environment compared to baseline levels of stereotypic behavior. Interestingly, Iwata et al. (1982/1994) stated that a behavior is determined to be automatically reinforced if it is high in the alone condition and low in the play condition, which is systematically enriched with preferred items and non-contingent attention. This statement suggests that it may be common for automatically reinforced behavior to be lower in conditions that have enriched environments. Piazza et al. expanded on previous research by exposing 3 participants to two different environmental enrichment conditions. One condition included hypothesized functionally matched stimuli and the other with hypothesized functionally unmatched stimuli. Results of this study showed that the percentage of 10-s intervals with stereotypic behavior was higher when the stimuli in the environment were unmatched compared to when they were matched (Piazza et al.).

Sensory extinction. Extinction as a procedure is characterized by the discontinuation of a contingency, which sometimes includes the elimination of some present reinforcer (Iwata, Pace, Cowdery, & Miltenberger, 1994). Iwata et al. evaluated procedural variations of extinction for treating problem behaviors. They exposed participants to both functionally indicated and nonindicated variations of extinction. They found that when the behavior no longer resulted in the specific reinforcer that maintained the behavior, the behavior decreased (Iwata et al.). Although Iwata et al. examined a variety of socially and non-socially

maintained problem behaviors, the data for the non-socially maintained responses showed that it is critical to discontinue the relationship between the response and the reinforcer for maximum treatment outcomes. Other studies have specifically examined the role of sensory extinction on problem behavior. Borrero, Vollmer, Wright, Lerman, & Kelley (2002) used protective equipment to break the contingency between SIB and the sensation it produces. The general purpose of the study was to assess the effects of protective equipment on functional analysis outcomes. Thus, the protective equipment was used because it allowed for the behavior to occur but broke the contingency between SIB and the sensation it naturally produces (i.e.; extinction). Results of this experiment showed that although protective equipment compromised interpretation of functional analyses, they suggested that extinction with the aid of protective equipment is a valid method to treat automatically reinforced SIB.

Differential reinforcement. Differential reinforcement, defined as reinforcing some responses in a response class while placing others on extinction, has also been used as a treatment for stereotypy (Rapp & Vollmer 2005). Although there are different types of differential reinforcement, including differential reinforcement of other (DRO), alternative (DRA), incompatible (DRI), and low or high rates of a behavior (DRL and DRH respectively), DRO and DRL are the most commonly used differential reinforcement operations to treat stereotypy (Cooper et al., 2007). Lanovaz et al. (2014) examined the effects of differential reinforcement of other behavior (DRO) on vocal stereotypy. They

found that a DRO procedure reduced vocal stereotypy immediately for two participants, in future sessions for a third participant, and had no effect on the level of vocal stereotypy for the fourth participant. Heffernan and Lyons (2016) used a DRO procedure to decrease an individual's automatically maintained SIB (i.e. nail biting). They found that the procedure was effective in reducing the SIB from over 50 instances per session to near zero rates, the DRO interval was thinned to a 60 min DRO. Singh, Dawson, & Manning (1981) delivered attention contingent on low rates of stereotypy or increased inter-response time (IRT) between stereotypic behaviors that was longer than the average IRT in baseline. The results of this procedure showed that stereotypy decreased and appropriate behavior increased during this DRL procedure.

Punishment. Punishment is often used to treat stereotypy (Rapp & Vollmer 2005). Punishment is defined as the presentation or removal of some stimulation following a behavior that functions to decrease that behavior under similar circumstances in the future (Cooper et al., 2007; Rapp & Vollmer). Punishment is often used after reinforcement-based treatments have been ineffective, and are often specifically used when the behavior is severe and requires immediate treatment (Fisher, Piazza, Bowman, Hagopian, & Langdon, 1994).

Punishment includes either an aversive stimulus presented contingent on a behavior (Type I, or positive punishment) or the removal of a preferred stimulus contingent on a behavior (Type II, or negative punishment; Cooper et al., 2007).

Negative punishment is probably more often used to treat problem behavior in homes and schools, as many common interventions are based on negative punishment (e.g., response cost and time out). For example, Zabel (1986) surveyed teachers and found that timeout procedures (also referred to as timeout from positive reinforcement; Donaldson, Vollmer, Yakich, & Van Camp, 2013) were used in 88% of preschools. Donaldson et al. evaluated a time-out procedure used in part to decrease the problem behavior of 6 participants during free play time. Results from their research found that the time-out effectively reduced the problem behavior of all 6 participants. They had an additional contingency in their study in which compliance with the instruction reduced the amount of time spent in time-out from 4 to 1 min, which increased compliance to the timeout instruction in 4 of the 6 participants.

Cole, Montgomery, Wilson, and Milan (2000) used another method of punishment (based on positive punishment) called positive practice overcorrection (PPOC) to treat self-stimulatory behavior. Overcorrection is a procedure in which an individual is required to engage in effortful behavior directly related to the problem behavior contingent on the occurrence of problem behavior. Results from their study showed that the occurrence of a target behavior was reduced effectively when the PPOC lasted 30 s, 2 min, or 8 min in duration. Cole et al. suggested that PPOC or overcorrection suppresses stereotypy, possibly as a function of the additional effort is required from the participant. It is also possible

that a decrease in stereotypy is a direct effect of physical contact being aversive in some way to the participant or learner (Rapp & Vollmer 2005).

Finally, some research has included empirical methods for determining what stimuli to use as punishers. Fisher et al. (1994) identified punishers for each of their two participants using a stimulus avoidance assessment. They found that when the aversive stimulus was delivered contingent on a target behavior, responding was reduced by 90% and 91.8% in their two participants. The stimulus avoidance assessment ensured that the stimuli chosen in upcoming sessions would function as punishers, increasing the likelihood of obtaining successful treatment effects.

Response Interruption and Redirection. Response Interruption and Redirection (RIRD) consists of a therapist introducing demands contingent on a behavior (Ahearn et al., 2007; Saini, Gregory, Uran, & Fanteiti, 2015). Saini et al. evaluated the effectiveness of using RIRD to treat stereotypic behaviors in 4 individuals. They extended research conducted by Ahearn et al. and that found that RIRD was effective for decreasing vocal stereotypy in children diagnosed with ASD (2007). Contingent on vocal stereotypy, the therapist presented vocal demands that interfered with the participant's ability to engage in the problem behavior. Therapists continued to issue demands until the participant complied with 3 consecutive demands (Ahearn et al.). Saini et al. extended this finding by examining the potential differential reduction of stereotypic behavior when participants were required to comply with 1 or 3 subsequent demands (RIRD 1

and RIRD 3 respectively). Results of Saini et al. showed that for all four participants, RIRD 1 was equally effective when compared to RIRD 3. Toper-Korkmaz et al. (2018) replicated Saini et al.'s evaluation of the effectiveness of RIRD 1 when compared to RIRD 3 and found corroborating evidence to support that these two RIRD methods are equally effective in reducing stereotypic behaviors.

It is clear that there are a multitude of potential treatments for automatically reinforced stereotypy, but it is not clear which treatment would be most effective for an individual. While there is little research on determining the most effective and efficient treatment procedure for automatically reinforced stereotypy, there is recent research on how to determine the most efficient treatment procedure for automatically reinforced SIB. Because automatically reinforced SIB and automatically reinforced stereotypy pose the same difficulties for treatment, it may be possible to adapt procedures that have been used to determine the best treatment for automatically reinforced SIB and use the same procedures to determine the best course of treatment for stereotypy.

Procedures to Determine Treatment of Automatically Maintained SIB

Prior to 2015, automatically reinforced behavior was generally considered to be a single functional category in the sense that (1) all non-socially mediated behavior produces its own reinforcement and (2) there was no accepted, systematic method for guiding treatment selection. Three recent studies evaluated procedures for (1) better understanding the nature of behaviors maintained by

automatic reinforcement and (2) increasing the chances of selecting successful treatments once it has been determined that behavior is maintained by automatic reinforcement. In a seminal study, Hagopian et al. (2015) developed a procedure designed to further categorize SIB maintained by automatic reinforcement, with the goal of creating a streamlined method of identifying a treatment that is most likely to be effective. Hagopian et al. took a 2-step approach to reaching this goal. First, they assessed whether a target behavior was maintained by automatic reinforcement, defined by the emergence of one of two patterns in the FA data: (1) high levels of responding in the alone condition and low in all other conditions or (2) high and variable levels of responding across all conditions (Iwata et al., 1982/1994; Iwata, et al. 1994). Next, they systematically assessed the level of differentiation between the play and alone conditions, suggesting that differential response patterns would indicate different and distinct functional properties of automatically reinforced SIB. They reviewed 39 cases of automatically reinforced SIB and categorized these cases into three categories using subtyping criteria based on patterns of responding in a functional analysis. Hagopian et al. (2017) applied these subtyping criteria to 49 additional data sets showing automatically maintained SIB. The purpose of this study was to replicate and test the generality of the findings from Hagopian et al. (2015). Additionally, Hagopian et al. (2017) assessed the efficacy of treatments designed to reduce SIB in each of the data sets collected to determine if subtype of the SIB indicated the effectiveness of different treatment procedures.

Finally, Berg et al. (2016) developed an integrated model for selecting interventions for responses maintained by automatic reinforcement. First, they conducted functional analyses of problem behavior, and enrolled participants whose results suggested that responding was maintained by automatic reinforcement. Next, Berg et al. compared levels of problem in a play versus alone/ignore condition to determine the extent to which access to alternative reinforcement competed with automatically reinforced behavior. Finally, the authors implemented treatments based on the results of the previous assessments.

Both studies by Hagopian and colleagues (2015, 2017) and Berg et al. (2016) endeavored to categorize automatically maintained behavior in such a way to improve treatment selection. In the next section, we describe the outcomes of those procedures.

Subtypes of Self-Injurious Behavior

Hagopian et al. (2015; 2017) identified two different response patterns in their FA data, both of which indicated that the SIB was maintained by automatic reinforcement: (a) SIB occurred at high levels in the alone or ignore condition but low levels in the play or control condition or (b) SIB occurred at high levels across both the alone or ignore condition and the play or control condition. Hagopian et al. categorized these patterns into Subtype 1 and Subtype 2, respectively (2015; 2017). Additionally, if the participant exhibited self-restraint behaviors in combination with SIB during the functional assessment, the behavior was categorized into a third Subtype (Subtype 3, Hagopian 2015; 2017).

The ultimate purpose of the subtyping procedures used by Hagopian et al. (2015, 2017) was to assess whether the categorization would be predictive of differential treatment efficacy. Hagopian et al. (2015) hypothesized that those in Subtype 1 would be more receptive to reinforcement-based treatment and that those in Subtype 2 would be more resistant to reinforcement-based treatment. They also hypothesized that the differences in patterns of responding between Subtype 1 and Subtype 2 were due to differences in the potency of the reinforcing consequences produced by SIB (Hagopian et al., 2015). Their results supported this hypothesis in that reinforcement-based procedures were effective in reducing SIB to acceptable levels for those in Subtype 1, but not Subtype 2 or 3 (Hagopian et al., 2015). In addition, they found a positive correlation between levels of differentiation in FA data between alone and play conditions and the percentage of reduction of SIB during reinforcement-based treatment (Hagopian et al., 2015, 2017). Results of Hagopian et al. 2017 reiterate the clinical value of identifying different subtypes of problem behavior by using the same subtyping procedure to subtype 49 additional data sets.

Subtyping can be thought of as a measure of sensitivity of the behavior to disruption (i.e. Subtype 1 is highly sensitive, while Subtype 2 is relatively insensitive; Hagopian et al., 2015; 2017). They found that those behaviors that are more sensitive to disruption require less invasive treatments while insensitive behaviors require more intensive treatments (Hagopian et al., 2015; 2017). Thus, assessing and categorizing automatically reinforced SIB allows prediction of the

most effective and efficient interventions for SIB in each category (Hagopian et al., 2015, 2017).

Berg et al. (2016) provided some support for Hagopian et al.'s (2015) study. Berg et al. delineated differential treatments for automatically maintained problem behavior based on the patterns of responding in the alone/ignore condition and the non-contingent reinforcement (NCR) play/control condition of a pairwise FA. Problem behavior that was differentiated across conditions, high in the alone/ignore and low in the NCR condition (roughly analogous to Hagopian et al.'s Subtype 1), was categorized as Pattern 1, and was treated with an NCR procedure. Problem behavior that was undifferentiated across conditions (roughly analogous to Hagopian et al.'s Subtype 2) was categorized as Pattern 2 or 3 contingent on the outcome of a concurrent-operants assessment. If a competing alternative stimulus was identified, the behavior was categorized as Pattern 2 and the behavior was treated using differential reinforcement with a response cost. If a competing alternative stimulus was not identified, the behavior was categorized as Pattern 3 and the behavior was treated using differential reinforcement with a response cost plus blocking. Results from this study show that patterns of responding in a pairwise FA on automatically maintained problem behavior may predict the success of treatments such as NCR, response cost, and blocking.

Purpose

Like SIB, there are several possible treatments for stereotypy, each with their own distinct advantages and disadvantages. Hagopian et al. (2015, 2017)

categorized automatically reinforced SIB into three subtypes, each with unique functional properties. They hypothesized and showed that these subtypes could be used to suggest different and specific intervention procedures that were most likely to be effective in decreasing SIB. Given the similarities between automatically reinforced SIB and stereotypy, it is possible that treatments for automatically reinforced stereotypy could be delineated using a similar subtyping procedure.

The purpose of the present research is to extend the Hagopian et al. (2015, 2017) categorization procedure to stereotypy, another behavior often maintained by automatic reinforcement. We subtyped stereotypic behaviors consistent with Hagopian et al. as (1) stereotypic behavior occurring at highest levels in the alone or ignore condition but low in the play or control condition is classified into Subtype 1 and (2) stereotypic behavior occurring at high and variable across both alone or ignore and play or control conditions is classified into Subtype 2 (2015; 2017). Next, we assessed the extent to which stereotypy categorized as Subtype 1 and Subtype 2 are differentially responsive to intervention. We hoped to replicate the findings of Hagopian et al. (2015, 2017) and following the categorization of stereotypy into Subtype 1 and Subtype 2, predict the most effective and efficient treatment to reduce the stereotypic behavior to clinically and developmentally acceptable levels.

General Methods

Participants and Settings

Participants in Study 1 (the pre-experimental assessment) included 6 individuals ranging in age from 2 to 9 who were receiving early intervention services, severe behavior services, and/or social skills services at an autism treatment center in Melbourne, Florida or in New Brunswick, New Jersey. Each participant had a previous diagnosis of ASD and engaged in visual, vocal, or motor stereotypy. All 6 individuals were exposed to functional analysis procedures that produced an automatic reinforcement function.

Participants in Studies 2 and 3 included an additional 3 individuals with a previous diagnosis of ASD who receive early intervention services at an autism treatment center in Melbourne, Florida. These three participants also displayed some topography of automatically maintained stereotypy based on criteria described in the literature (e.g., Hagopian et al., 2017; Iwata et al., 1982/1994; Querim et al., 2013; Vollmer et al., 1995).

Response Measurement and Interobserver Agreement (IOA)

Each topography of stereotypic behavior for individual participants was operationally defined based on the body part engaged in the stereotypic movement (Hagopian et al., 2015). These definitions were tailored to the individual, and based on the participants' individual clinical Behavior Intervention Plans (BIP), if applicable. To be included as participants, caregivers and the clinical team must have reported that the behaviors generally persisted

over time, occur across environments, and were out of synchrony with typical development (Rapp & Vollmer 2005).

To calculate interobserver agreement (IOA) for Studies 2 and 3, both observers coded a minimum of 20% of the data sets to evaluate treatment integrity. The extent to which observers agreed was calculated using 10-s interval-by-interval IOA by counting the total number of intervals in which the observers either agreed that stereotypy occurred or agreed that it did not occur. This number was then divided by the total number of 10-s intervals in a 5-min session. The resulting number was then multiplied by 100 to convert it into a percentage. The average IOA for each participant and all sessions must have been above 80%. This includes the competing items assessment, the FA, and the treatment for all five topographies of stereotypy.

Interobserver agreement for Nathan was taken for 75% of sessions in the competing items assessment and 76.92% of sessions in the FA for both his vocal and motor stereotypy, and 60% and 55%, respectively, for his vocal and motor stereotypy in the treatment phase. In the competing items assessment, the average agreement was 83.03% across both topographies of stereotypy and item engagement. In the FA, the average agreement was 88% for his vocal stereotypy, and 85.67%, for his motor stereotypy. In the treatment phase, the average agreement was 86.39% for his vocal stereotypy, and 85.67% for his motor stereotypy.

Interobserver agreement for Seth was taken for 62.5% of sessions in the competing items assessment, 68.75% of sessions in the FA, and 87.5% of sessions in the treatment phase for his visual stereotypy. In the competing items assessment, the average agreement was 88.31% for his visual stereotypy and item engagement. In the FA, the average agreement was 91.21% for his visual stereotypy. In the treatment phase, the average agreement was 85.71% for his visual stereotypy.

Interobserver agreement for Miguel was taken for 70% of sessions in the competing items assessment and 53.85% of sessions in the FA for both his vocal and visual stereotypy, and 45.45% and 34.48%, respectively, for his vocal and visual stereotypy in the treatment phase. In the competing items assessment, the average agreement was 86.39% across both topographies of stereotypy and item engagement. In the FA the average agreement was 96.19% for his vocal stereotypy and 86.19% for his visual stereotypy. In the treatment phase, the average agreement was 93% for his vocal stereotypy and 88.33% for his visual stereotypy.

Study 1: Pre-Experimental Assessment

Rationale and Purpose

The purpose of Study 1 was to extend the categorization procedure (Hagopian et al., 2015, 2017) to stereotypy, an alternative topography that is often maintained by automatic reinforcement, and to set the occasion for Studies 2 and

3, in which we tested the efficacy of the subtyping procedure for treatment selection.

Participants

Participants included 6 individuals whose functional analysis results suggested that stereotypic behavior was maintained by automatic reinforcement. These individuals included 5 males and 1 female (ages 3 to 9) who engaged in at least one topography of stereotypy. Five participants engaged in only one topography of stereotypy (vocal, motor, or hair pulling) and one participant engaged in two topographies (vocal and motor).

Procedure

We applied criteria used by Hagopian et al. (2015) to determine whether pre-existing functional analyses of stereotypic behaviors could be subtyped based on response patterns. To subtype an FA data set, the experimenter drew two criterion lines on the data set collected in the functional analysis (see Figure 1). The upper criterion line (UCL) was drawn between the second and third highest data points and the lower criterion line (LCL) was drawn between the second and third lowest data points in the play condition. Then, the number of data points in the alone or ignore condition that fall below the LCL were subtracted from the number of data points that fell above the UCL. The difference was then divided by the total number of data points in the alone or ignore condition. This numerical value is referred to as the quotient score (Hagopian et al., 2015, 2017). In this study, the lines were drawn between the first and second highest and lowest data

points due to there only being 3 data points in several of the play conditions of the data analyzed. After calculating the quotient score, data sets were then categorized into Subtype 1 and Subtype 2 based on whether their quotient score was higher or lower than 0.5, respectively.

Criteria for Subtype 1. Stereotypy was categorized as Subtype 1 if the FA showed a clear differentiation in the level of stereotypy between the alone or ignore conditions and the play condition. More specifically, the level of stereotypy in these data sets are high in the alone/ignore condition but low in the play or control condition. Those who met criteria for Subtype 1 have a quotient score greater than or equal to 0.5.

Criteria for Subtype 2. Stereotypy was categorized as Subtype 2 if the FA showed little to no differentiation in the level of stereotypy between the alone or ignore conditions and the play condition. More specifically, the level of stereotypy in these data sets are high (and sometimes variable) across both the alone or ignore condition and the play or control condition. Specifically, either (a) the quotient score for these FAs was less than 0.5, (b) over 30% of data points were overlapping between the two conditions, or (c) the percentage of time spent engaged in stereotypic behaviors was more than 40% in all conditions. This percentage was chosen based on the range of time allocated to stereotypy provided by Rapp and Vollmer (2005).

Results and Discussion

Six participants produced 7 data sets, as 1 participant engaged in multiple topographies of stereotypy. Four data sets met the criteria for Subtype 1 (see Figure 2) and three met the criteria for Subtype 2 (see Figure 3). Interestingly, for the one participant who engaged in both vocal and motor stereotypy, vocal stereotypy met the criteria for Subtype 1 while motor stereotypy met the criteria for Subtype 2. These results emphasize the importance of separating various topographies of behavior into subtypes, even if the topographies are maintained by the same general reinforcement contingency (i.e., automatic reinforcement). In general, these results demonstrate that the categorization methods designed by Hagopian et al. (2015) can be used to categorize automatically reinforced stereotypy.

Study 2: Prospective Functional Analyses and Subtyping

Rationale and Purpose

The purposes of Study 2 were to (1) prospectively test the generality of the Hagopian et al. (2015, 2017) subtyping procedures and (2) identify participants for Study 3.

Participants and Setting

Participants included three 4-year old individuals diagnosed with ASD who were reported to engage in stereotypy and were enrolled in early intervention services at an autism treatment center in Melbourne, Florida. Nathan was a 4-year-old boy who engaged in *vocal* and *motor* stereotypy. His *vocal stereotypy*

was defined as any vocal behavior that was nonfunctional, repetitive, and also included repetitive blowing. His *motor stereotypy* was defined as moving his fingers or objects back and forth between 2.54 cm and 20.3 cm away from his face. As his motor stereotypy in the clinical setting often involved items such as stick-shaped items (i.e. drumsticks, markers, pencils, etc.), he was given access to a red smooth drumstick and a blue ridged drumstick in all conditions. This was to ensure that his motor stereotypic behavior was available in all conditions to get an accurate representation of the effects of both the assessment procedures as well as the treatment effects. Seth was a 4-year-old boy who engaged in *visual* stereotypy. His *visual stereotypy* was defined staring at the walls of the room and rolling his eyes, using his fingers to make shadows on the floor and staring at them, spinning in a circle while staring at the ceiling, and looking at his hands while he moved them between 2.54 and 20.3 cm from his face. Additionally, pressing his face to the walls of the room and staring at the walls was also considered visual stereotypy. Miguel was a 4-year-old boy who engaged in *vocal* and *visual* stereotypy. His *vocal stereotypy* was defined as any vocal behavior that was nonfunctional, repetitive, and also included repetitive counting or spelling. His *visual stereotypy* was defined as holding his fingers within 20.3 cm of his face, often counting on them or spelling in American Sign Language. Additionally, drawing numbers and letters on the walls of the room with his fingers was considered visual stereotypy.

All sessions across all phases of assessment conditions and treatment were conducted in a 3.35 m x 3.5 m treatment room with the same therapist in the room with the participant. This room was a barren environment, however a single chair was present for the therapist to sit in while sessions were being conducted. This environment was the same across all sessions and individuals, except Nathan who was also given access to two sticks during all sessions.

General Procedures

Experimenters exposed the participants to a variety of assessments. Figure 4 shows the potential flow through the three assessments and experimental condition. As seen in Figure 4, all participants experienced a preference assessment, a competing items assessment, and an adapted pairwise functional analysis (Fisher et al., 1992; 2004).

Preference assessment (PA). Experimenters conducted a multiple stimulus without replacement preference assessment (DeLeon & Iwata, 1996) to identify a hierarchy of preferred stimuli to use in the subsequent competing items assessment. Staff and caregivers nominated preferred items for use in the assessment. Experimenters conducted a series of trials in which three items were presented to the participant in a line on the floor approximately 15 cm apart. The participant was instructed to choose one. A selection was defined as the item the participant touched first following the instruction to make a choice. After a selection was made, the participant was given 20 s of free access to the selected item. Following this access period, the item was removed from the array and the

remaining items were rearranged and again presented. If the participant attempted to select multiple items simultaneously, the therapist blocked the attempt and represented the stimuli and the instruction to “choose one”. If a participant did not select an item, the therapist prompted the participant to sample each item for 5 s. Next, the therapist re-presented the stimuli and the instruction “choose one”. If the participant did not choose an item again, the remaining items were recorded as “not selected”. This therapist conducted this assessment up to three times for each participant to increase the probability of acquiring an accurate hierarchy of each individual’s preference.

Competing items (stimulus) assessment. The therapist completed a competing items assessment for all participants following the MSWO. The therapist selected stimuli to be included in the assessment for each participant based on verbal report from each participant’s clinical team. Each item was presented two times for 10 min each. During each session, observers recorded item engagement and the occurrence of stereotypy using momentary time sampling with 6-s intervals. The number intervals in which the individual was engaged with the item at the end of the 6-s interval was divided by the total number of intervals. This was then converted into a percentage of total intervals in which the individual was engaged with that item. Additionally, the percentage of total intervals in which the individual was engaged with an item and stereotypic behaviors was calculated for the first 2 min, the first 5 min, and the last 5 min. The item that best competed with stereotypy, defined as yielding the greatest

reduction in the percentage of intervals in which the individual was engaged in stereotypic behaviors as compared to baseline, was then used in the FA and subsequent environmental enrichment portion of the treatment phase.

Functional analysis. Experimenters exposed participants to an adaptation of Querim et al.'s (2013) automatic reinforcement pre-screening methodology. Querim et al. conducted a series of alone/no interaction conditions to determine whether the behaviors maintained by social or automatic reinforcement. Following an extended ignore condition, we conducted a pairwise functional analysis that include (a) ignore conditions and (b) play conditions. This pairwise FA allowed us to (a) determine if the behavior is automatically maintained and (b) categorize the behavior into either Subtype 1 or Subtype 2. Behavior was determined to be maintained by automatic reinforcement if responding persisted across conditions, or if the responding was differentiated between conditions and was high in the ignore condition and low during the play condition (Iwata et al., 1982/1994). These different patterns of responding allowed us to categorize the stereotypy into subtypes, as is the purpose of this study. Further, if responding was suppressed during the play condition, we concluded that the stimuli competed with the maintaining reinforcer for stereotypy (Berg et al., 2016). We hypothesized that (a) those with suppressed stereotypic behaviors in the play conditions will be more receptive to less intrusive treatments such as environmental enrichment, and (b) those whose stereotypic behaviors persist

across conditions will require more intrusive treatments, such as punishment based treatments to treat stereotypy.

Sessions were counterbalanced in an ABBABAAB format (Barlow & Hayes 1979). Sessions occurred in a treatment room (3.35 m x 3.5 m), and lasted 5 min. A single therapist was in the room with the participant during each condition and they were provided with a chair to sit in. In the *ignore* condition, there were no programmed consequences for engaging in stereotypy. *Play* conditions were used as a control condition. In the *play* condition, participants had free access to the item that best competed with their stereotypic behavior as was identified in the previous competing items assessment. In the play condition, the therapist also provided attention on a 30-s fixed-time schedule (Iwata et al., 1982/1994).

Subtype criteria and coding. Pairwise functional analysis data were then analyzed in a manner identical to that described by Hagopian et al. (2015, 2017) except we analyzed stereotypic behavior as opposed to self-injurious behavior.

Results

Figure 5 shows data for the 3 participants' competing items assessments. These data show both the percentage of time each participant spent engaged with the item and spent engaged in stereotypic behavior. For Nathan, both topographies met criteria for Subtype 1. In the competing items assessment 2 of the 3 stimuli tested effectively competed with his motor stereotypy, while only 1 of the 3 stimuli tested effectively competed with his vocal stereotypy. The only

item that best competed with both topographies of stereotypy was the iPad. For Seth, his visual stereotypy met criteria for Subtype 1. In the competing items assessment, 3 of the 3 stimuli tested competed with his visual stereotypy to some extent. However, the item that best competed with this topography of stereotypy was the iPad. For Miguel, both topographies met criteria for Subtype 1. In the competing items assessment, 3 of the 3 stimuli tested effectively competed with both his visual and vocal stereotypy. However, the item that best competed with both topographies of his stereotypy was the iPad.

While the iPad was the item that best competed with each individual's stereotypic behaviors, this item was presented slightly differently for each participant based on their ability to utilize this device. For Nathan, the in-session researcher held the device and the device played Nathan's favorite video on repeat. For Seth, the in session researcher opened the YouTube Kids® app on the device, and ensured that this app was constantly open. For example, if Seth went back to the home screen the therapist reopened the app. For Miguel, the iPad was simply available for him to choose to engage with it in any way he wished (e.g. watch videos, play with the timer, play a game on the device, etc.).

Figure 6 shows the 5 functional analyses from the 3 participants. Each data set collected shows clear differentiation between the test and control conditions, and met the criteria for Subtype 1 (quotient score above 0.5). Of the 3 participants that went through this study, none of their stereotypy showed that they fit the criteria for Subtype 2 (quotient score below 0.5). Nathan's motor

stereotypy received a quotient score of .75 while his vocal stereotypy received a score of 1. Seth's visual stereotypy and both Miguel's vocal and visual stereotypy also all received a score of 1.

Study 3: Treatment Analyses

Participants and Setting

Participants included the 3 individuals from Study 3. Sessions were conducted in 3.35 m x 3.5 m treatment rooms. These rooms were equipped with one chair for the therapist in the room to sit in, as was consistent with Study

General Procedures

Following Study 2, experimenters exposed the participants to a series of interventions progressing from least-to-most intrusive, defined by the extent to which the therapist interacts with the participant and actively interrupts stereotypic behavior. Figure 4 shows the flow of treatments as follows: Enriched Environment (EE), punisher assessment, contingent punishment. Once a treatment was deemed effective (defined by a clear differentiation between the baseline and treatment conditions), the participant no longer progressed through the other treatments. Ultimately, the number of treatment procedures a participant experienced depended on the behavioral patterns that emerged over the course of the treatment progression (Berg et al., 2016).

Enriched environment (EE). Experimenters initially exposed all participants to Enriched Environment (EE) following the pairwise functional analysis. During environmental enrichment conditions, the stimuli identified as

most likely to reduce stereotypy in the competing items assessment were available upon the start of session (Watkins & Rapp, 2014). Additionally, participants were given attention on a FT-30 s schedule as was delivered in the play condition of the FA. It is hypothesized that only those who are categorized into Subtype 1 will show acceptable decreases in the level of stereotypic behavior during this treatment condition (Hagopian et al., 2015; 2017). Because this treatment method is essentially the same as the play condition in the FA, this procedure is likely to be effective in treating the stereotypy of those categorized into Subtype 1.

Results

Figure 7 shows treatment data for each of the 3 participants. Treatment was arranged in a reversal design (ABAB) to demonstrate experimental control of the treatment component. Each of these three participants, for all five behaviors measured across, met the criteria for Subtype 1. As was hypothesized, those topographies that met criteria for Subtype 1 were responsive to reinforcement-based treatments. The data supports this as each topography of stereotypy was reduced through environmental enrichment.

Discussion

In the current series of studies, we evaluated the extent to which the method of subtyping designed by Hagopian et al. (2015) used to subtype automatically maintained SIB could be used to subtype stereotypy. Further, we evaluated the hypothesis first postulated by Hagopian et al. (2017) suggesting that those with behaviors classified as Subtype 1 would require less intrusive

reinforcement-based treatments while those classified as Subtype 2 would require more intrusive treatments. We applied and extended this procedure to the treatment of stereotypy. The ultimate applied purpose of this study was to mitigate the use of unnecessary punishment for stereotypy as it is a non-dangerous behavior.

In the first study, we replicated the retrospective subtyping procedures developed by Hagopian et al. (2015). Hagopian posited that automatically maintained self-injurious behavior could be categorized into three subtypes based on behavioral patterns displayed in a functional analysis. Our study extended this concept of categorizing behavior to stereotypy and found that stereotypy can be categorized into two subtypes based on criteria defined by Hagopian et al. (2015).

Study 2 was specifically designed to test the hypothesis that the subtyping procedure would be useful for prospectively differentiating between two specific kinds of stereotypy: one that appears sensitive to disruption with environmental enrichment, and one that does not. Results of show that automatically reinforced stereotypy can be categorized into two subtypes based on the patterns of responding in an FA. However, future researchers might look to extend this subtyping procedure to other topographies of behavior, specifically to behaviors that are often automatically maintained.

Following subtyping, Study 3 was conducted to determine if the subtype of behavior was predictive of the success of a specific treatment. Experimenters planned to expose participants to up to three treatment procedures to determine a

relationship between the subtype and an effective treatment. These treatment procedures were arranged to progress from least intrusive to most intrusive. We hypothesized that those whose behaviors fell into Subtype 1 would be more sensitive to reinforcement based treatment and that those whose behaviors fell into Subtype 2 would require more intrusive treatments (e.g., punishment based treatments) to reduce their levels of stereotypy. However, we found no data sets that would have been classified as Subtype 2 and were therefore unable to support or refute this hypothesis.

Three out of three individuals, and five out of five topographies of stereotypy, were categorized as Subtype 1 and were responsive to an environmental enrichment condition. If individuals whose stereotypy was classified as Subtype 2 required punishment-based procedures, this would have enhanced the predictive power of the subtyping procedure for an individual's stereotypy. However, we were unable to fully test this hypothesis as we did not identify any participants whose stereotypy was classified as Subtype 2, nor did any of our participants require punishment-based procedures to reduce their stereotypy. Had we found individuals whose behavior was categorized into Subtype 2, we would have conducted a stimulus avoidance assessment, a punisher assessment, and determined a punishment-based treatment procedure for reducing stereotypy (Fisher et al., 1994). As we did not find individuals who met criteria to be categorized into Subtype 2, more research is needed to determine how those with Subtype 2 stereotypy responds to treatment.

Not identifying individuals who met criteria for Subtype 2 was both a limitation of Study 3 and was surprising for at least two reasons. First, the retrospective subtyping procedure used in Study 1 yielded 7 data sets, of which 3 were categorized as Subtype 2. Based on the results of Study 1, we expected that 1-2 out of the 5 data sets would meet Subtype 2 criteria. Second, our participants' clinical teams reported that these individuals engaged in stereotypy for large portions of session time. It was further unexpected that we found at least 1 item that competed with both topographies of stereotypy for each participant in the competing items assessment. Future researchers should continue to use this subtyping procedure to find individuals whose stereotypic behaviors are categorized into Subtype 2. Additionally, future researchers could work to improve upon this subtyping procedure to better categorize stereotypic behaviors into these subtypes.

One reason we may not have found any topographies of stereotypy in this study that met the requirements for Subtype 2 may be due to the competing items assessment that was ran as part of Study 2. This items assessment was used to determine which of three items would compete with and reduce the level of the participants' stereotypy most effectively. Only the item that decreased the individual's stereotypy to the greatest degree was used in play condition of the FA. Thus, it could be argued that the play condition was engineered to reduce stereotypy more than the play condition of a typical FA, which often uses the individual's most preferred items. This may explain why we found stereotypy that

was categorized into Subtype 2 in Study 1 and not in Study 2. The participants used in Study 1 were exposed to play conditions in which they had access to their most preferred items rather than items that were likely to compete with their stereotypy, which may not always be the individual's most preferred item.

This could be considered a limitation of this study in that the play conditions of the FA that were used in Study 1 as compared to Study 2 were slightly different, which may have contributed to not finding topographies that were classified as Subtype 2. However, this limitation is a potential topic of future study. Future researchers could compare the results of functional analyses that use play conditions with competing items compared to play conditions with preferred items. Further, future researchers may compare the results of different types of functional analyses (e.g., full versus a pairwise, or versus a pairwise versus a trial-based FA) to determine if the same behavior yields different subtypes when different functional analyses are used.

In summary, results from this research indicate that the subtype into which the behavior falls is in part predictive of the efficacy of treatments such as environmental enrichment. Due to the predictive power of the Subtyping procedure, it is possible that after conducting a pairwise adaptation of the Quirem et al. (2013) FA and subtyping the behavior, the most likely to be effective treatment could be selected. This would decrease the amount of unnecessary punishment procedures used to treat stereotypy. This procedure described in this

research provides a possible method of further classifying stereotypic behaviors and delineated a most likely to be effective treatment.

References

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5™* (5th ed.). Arlington, VA: American Psychiatric Publishing, Inc. http://dx.doi.org/10.1176/appi.books.9780890425596

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis,12*(2), 199-210. doi:10.1901/jaba.1979.12-199

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1–21. http://dx.doi.org/10.1002/jaba.30

Berg, W. K., Wacker, D. P., Ringdahl, J. E., Stricker, J., Vinquist, K., Dutt, A. S. K., Mews, J. (2016). An integrated model for guiding the selection of treatment components for problem behavior maintained by automatic reinforcement. *Journal of Applied Behavior Analysis, 49*, 617–638. http://dx.doi.org/10.1002/jaba.303

Borrero, J. C., Vollmer, T. R., Wright, C. S., Lerman, D. C., & Kelley, M. E. (2002). Further evaluation of the role of protective equipment in the functional analysis of self-injurious behavior. *Journal of Applied Behavior Analysis, 35*, 69–72. http://dx.doi.org/10.1901/jaba.2002.35-69

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18*, 111–126. http://dx.doi.org/10.1901/jaba.1985.18-111

Celiberti, D. A., Bobo, H. E., Kelly, K. S., Harris, S. L., & Handleman, J. S. (1997). The differential and temporal effects of antecedent exercise on the self-stimulatory behavior of a child with autism. *Research in Developmental Disabilities, 18*, 139–150. http://dx.doi.org/10.1016/S0891-4222(96)00032-7

Cole, G. A., Montgomery, R. W., Wilson, K. M., & Milan, M. A. (2000). Parametric analysis of overcorrection duration effects: Is longer really better than shorter? *Behavior Modification, 24*, 359–378. http://dx.doi.org/10.1177/0145445500243004

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Applied behavior analysis.

Conners, J., Iwata, B. A., Kahng, S., Hanley, G. P., Worsdell, A. S., & Thompson, R. H. (2000). Differential responding in the presence and absence of discriminative stimuli during multielement functional analyses. *Journal of Applied Behavior Analysis, 33*, 299–308. http://dx.doi.org/10.1901/jaba.2000.33-299

Deleon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis,29*(4), 519-533. doi:10.1901/jaba.1996.29-519

Donaldson, J. M., Vollmer, T. R., Yakich, T. M., & Van Camp, C. (2013). Effects of a reduced time-out interval on compliance with the time-out instruction. *Journal of Applied Behavior Analysis, 46*, 369–378. http://dx.doi.org/10.1002/jaba.40

Dunlap, G., Dyer, K., & Koegel, R. L. (1983). Autistic self-stimulation and intertrial interval duration. *American Journal of Mental Deficiency, 88*, 194–202.

Fisher, W. W., DeLeon, I. G., Rodriguez-Catter, V., & Keeney, K. M. (2004). Enhancing the effects of extinction on attention-maintained behavior through noncontingent delivery of attention or stimuli identified via a competing stimulus assessment. *Journal of applied behavior analysis*, *37*(2), 171-84.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., & Langdon, N. A. (1994). Empirically derived consequences: A data-based method for prescribing treatments for destructive behavior. *Research in Developmental Disabilities, 15*, 133–149. http://dx.doi.org/10.1016/0891-4222(94)90018-3

Fisher, W. W., Piazza, C. C., & Chiang, C. L. (1996). Effects of equal and unequal reinforcer duration during functional analysis. *Journal of Applied Behavior Analysis, 29*, 117–120. http://dx.doi.org/10.1901/jaba.1996.29-117

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*, 523–543. http://dx.doi.org/10.1002/jaba.236

Hagopian, L. P., Rooker, G. W., Zarcone, J. R., Bonner, A. C., & Arevalo, A. R. (2017). Further analysis of subtypes of automatically reinforced sib: A replication and quantitative analysis of published datasets. *Journal of Applied Behavior Analysis, 50*, 48–66. http://dx.doi.org/10.1002/jaba.368

Hammond, J. L., Iwata, B. A., Rooker, G. W., Fritz, J. N., & Bloom, S. E. (2013). Effects of fixed versus random condition sequencing during multielement functional analyses. *Journal of Applied Behavior Analysis, 46*, 22–30. http://dx.doi.org/10.1002/jaba.7

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185. http://dx.doi.org/10.1901/jaba.2003.36-147

Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function-based interventions. *Journal of Applied Behavior Analysis, 38*, 51–65. http://dx.doi.org/10.1901/jaba.2005.6-04

Heffernan, L., & Lyons, D. (2016). Differential Reinforcement of Other Behaviour for the Reduction of Severe Nail Biting. *Behavior Analysis in Practice, 9* (3), 253-256. doi:10.1007/s40617-016-0106-3

Horner, R. D. (1980). The effects of an environmental enrichment program on the behavior of institutionalized profoundly retarded children. *Journal of Applied Behavior Analysis, 13*, 473–491. http://dx.doi.org/10.1901/jaba.1980.13-473

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982/1994). Toward a functional analysis of self-injury. *Analysis & Intervention in Developmental Disabilities, 2*, 3–20. http://dx.doi.org/10.1016/0270-4684(82)90003-9

Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis,27*(1), 131-144. doi:10.1901/jaba.1994.27-131

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis, 27*, 215–240. http://dx.doi.org/10.1901/jaba.1994.27-215

Kern, L., Koegel, R. L., & Dunlap, G. (1984). The influence of vigorous versus mild exercise on autistic stereotyped behaviors. *Journal of Autism and Developmental Disorders, 14*, 57–67. http://dx.doi.org/10.1007/BF02408555

LeBlanc, L. A., Patel, M. R., & Carr, J. E. (2000). Recent advances in the assessment of aberrant behavior maintained by automatic reinforcement in individuals with developmental disabilities. *Journal of Behavior Therapy and Experimental Psychiatry, 31*, 137–154. http://dx.doi.org/10.1016/S0005-7916(00)00017-3

MacDonald, R., Green, G., Mansfield, R., Geckeler, A., Gardenier, N., Anderson, J., Sanchez, J. (2007). Stereotypy in young children with autism and typically developing children. *Research in Developmental Disabilities, 28*, 266–277. http://dx.doi.org/10.1016/j.ridd.2006.01.004

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*, 509–522. http://dx.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 of the effects of matched stimuli on behaviors maintained by automatic reinforcement. *Journal of Applied Behavior Analysis, 33*, 13–27. http://dx.doi.org/10.1901/jaba.2000.33-13

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*, 47–60. http://dx.doi.org/10.1002/jaba.26

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

Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The Behavior Problems Inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders, 31*, 577–588. http://dx.doi.org/10.1023/A:1013299028321

Saini, V., Gregory, M. K., Uran, K. J., & Fantetti, M. A. (2015). Parametric analysis of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis, 48*, 96–106. http://dx.doi.org/10.1002/jaba.186

Singh, N. N., Dawson, M. J., & Manning, P. (1981). Effects of spaced responding DRL on the stereotyped behavior of profoundly retarded persons. *Journal of Applied Behavior Analysis, 14*, 521–526. http://dx.doi.org/10.1901/jaba.1981.14-521

Thomason-Sassi, J. L., Iwata, B. A., Neidert, P. L., & Roscoe, E. M. (2011). Response latency as an index of response strength during functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 44*, 51–67. http://dx.doi.org/10.1901/jaba.2011.44-51

Thompson, R. H., & Iwata, B. A. (2007). A comparison of outcomes from descriptive and functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 40*, 333–338. http://dx.doi.org/10.1901/jaba.2007.56-06

Toper-Korkmaz, O., Lerman, D. C., & Tsami, L. (2018). Effects of toy removal and number of demands on vocal stereotypy during response interruption and redirection. *Journal of Applied Behavior Analysis*. Advance online publication. http://dx.doi.org/10.1002/jaba.497

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

Vollmer, T. R., Iwata, B. A., Duncan, B. A., & Lerman, D. C. (1993). Extensions of multielement functional analyses using reversal-type designs. *Journal of Developmental and Physical Disabilities, 5*, 311–325. http://dx.doi.org/10.1007/BF01046388

Zabel, M. K. (1986). Timeout use with behaviorally disordered students. *Behavioral Disorders, 12*, 15–21.

Figure 1

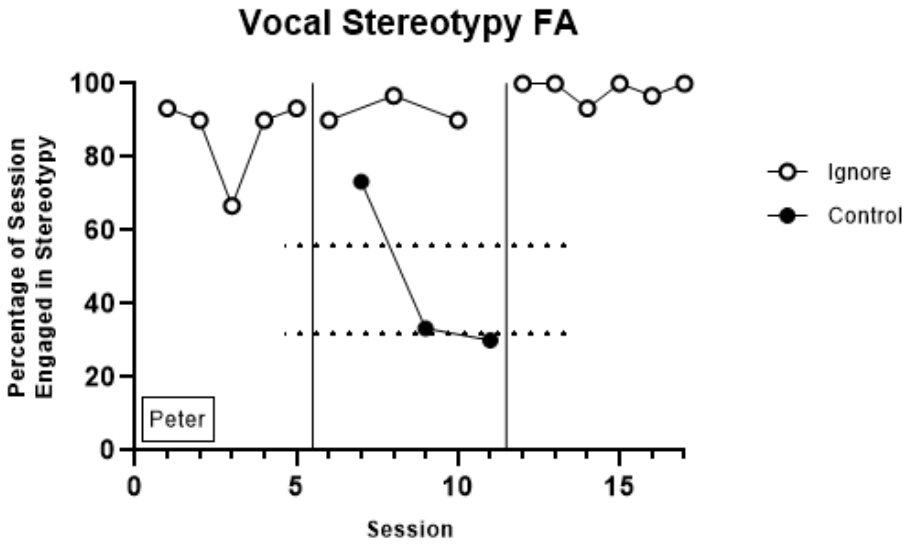

Figure 1: This figure shows an example of a graph with the Upper and Lower Criterion Lines (UCL and LCL respectively) drawn on it to determine the equation used to identify the quotient score for this individual. *Equation:* (3 - 0) / 3 = 1; Quotient Score = 1 therfore this behavior is classified into Subtype 1.

Figure 2

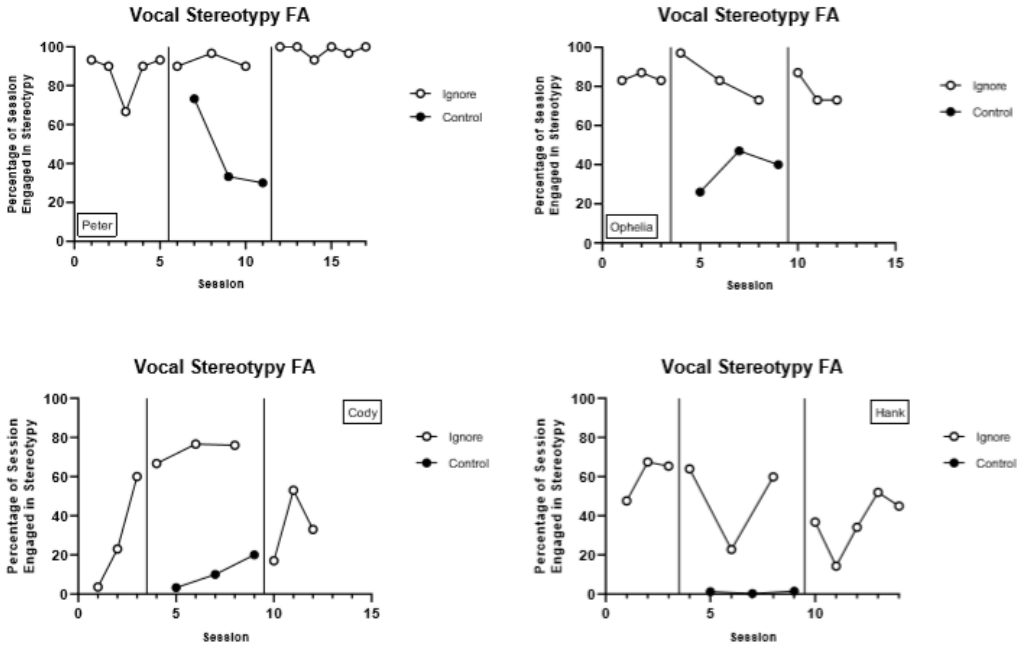

Figure 2: This figure shows stereotypic behavior categorized into Subtype 1, due to differentiation between conditions of the FA. The quotient score was calculated for each of these data sets and the scores for each of these data sets fell above 0.5.

Figure 3

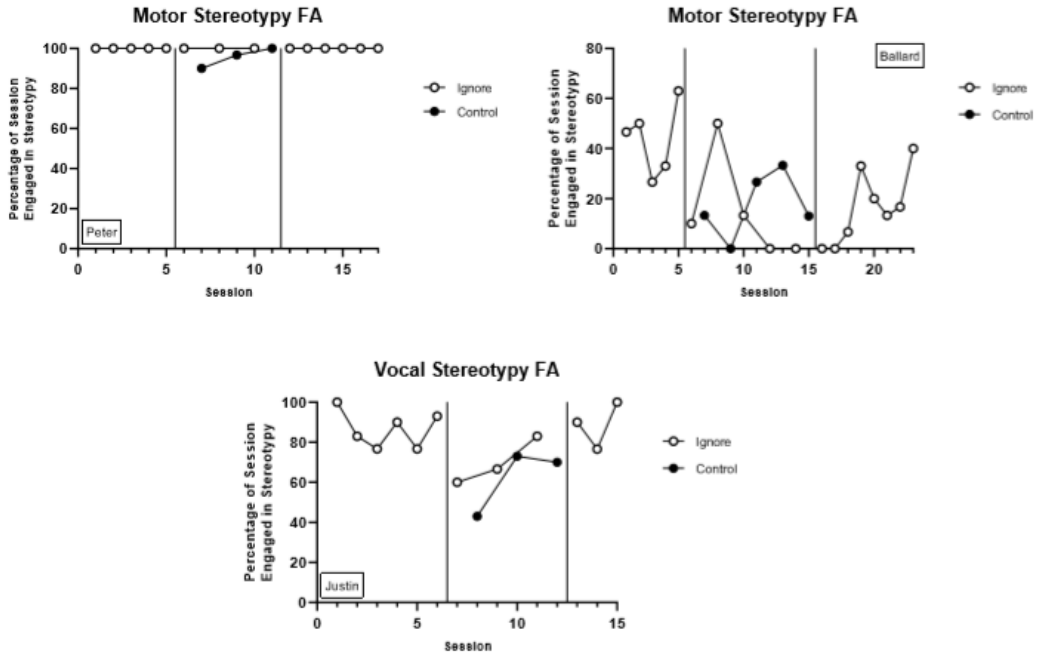

Figure 3: This figure shows stereotypic behavior categorized into Subtype 2, due to either no differentiation between levels of stereotypy between conditions of the FA, or the stereotypy occurred for over 40% of all sessions. The quotient score was calculated for each of these data sets and the scores for each of these data sets fell below 0.5.

Figure 4

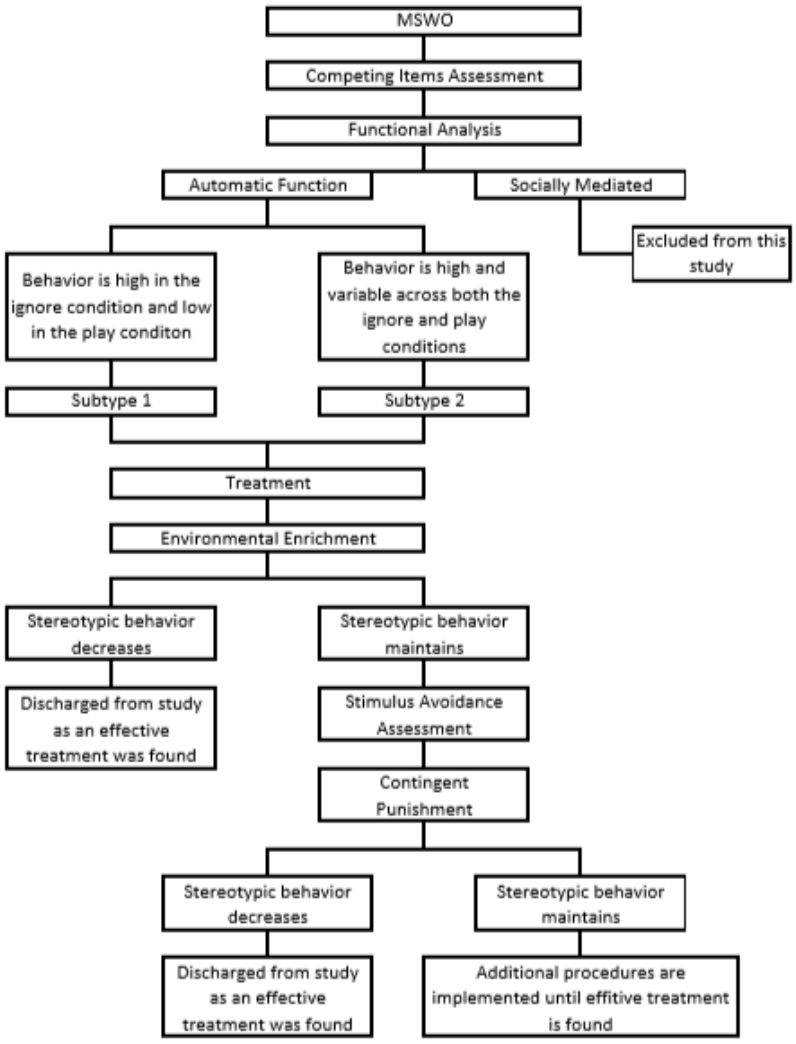

Figure 4. This figure shows the order of assessments each participant in Study 2 and Study 3 will experience as well as the treatment progression they will each experience. Once an effectively reduces stereotypy, the experimenter will terminate the treatment progression and the participant will be discharged from the study.

Figure 5

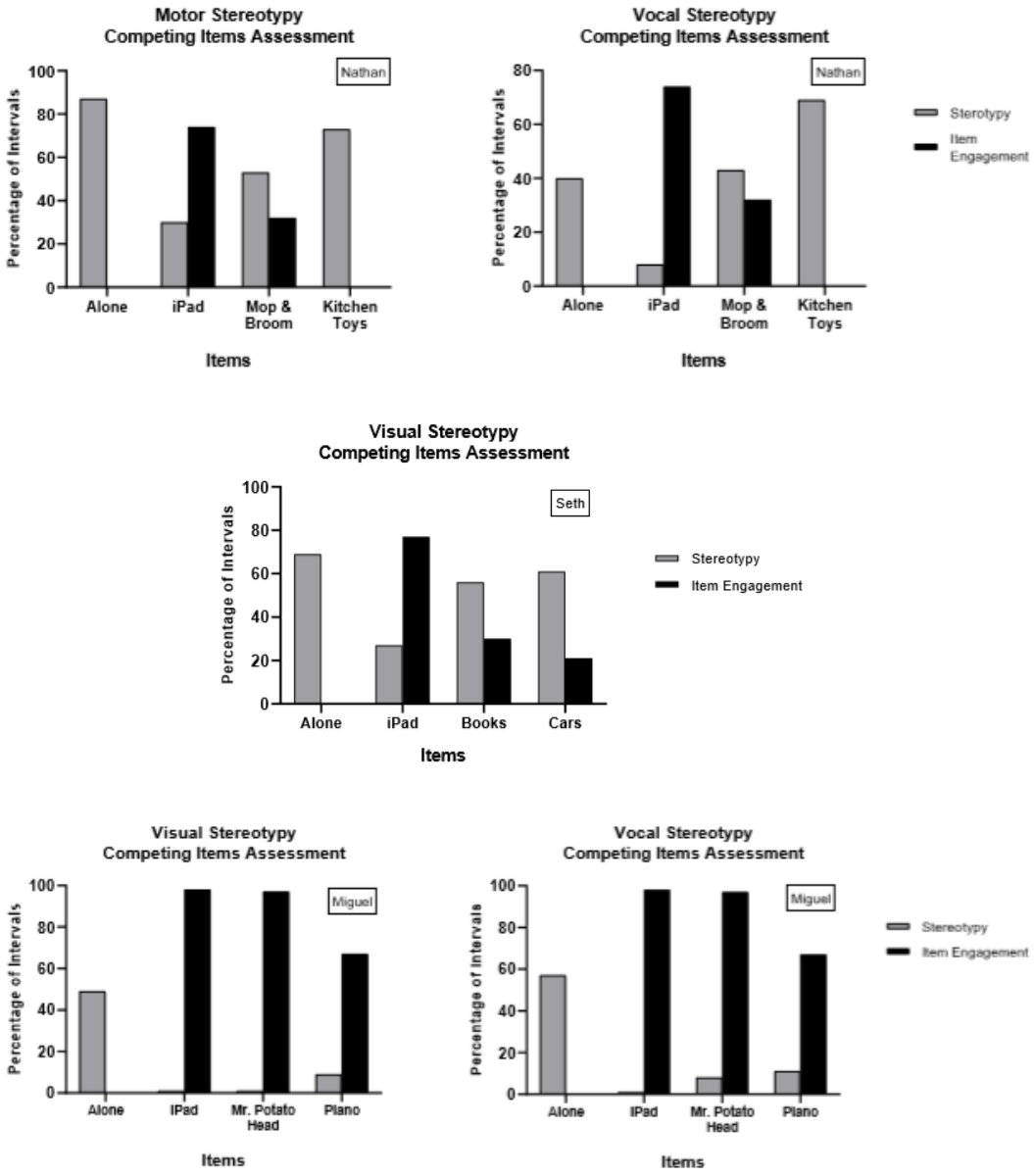

Figure 5. Figure 5 shows the results of a competing items assessment which was conducted to assess to what extend three stimuli compete with the stereotypy exhibited by each participant. The iPad reduced five out of five topographies of stereotypy for all three of the three participants.

Figure 6

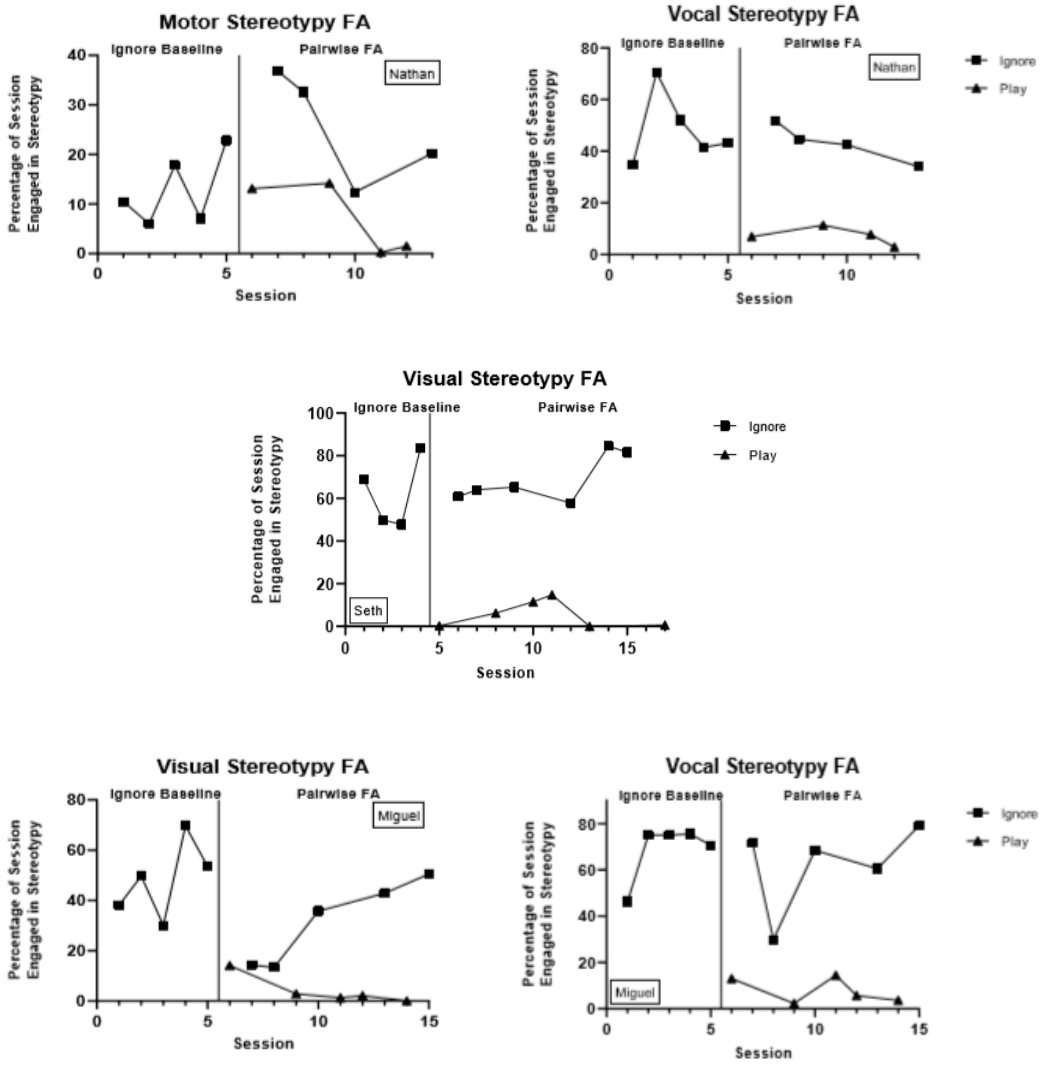

Figure 6. Figure 6 shows FA data from 3 participants that will be used for subtyping procedures. Each of the five data sets show clear differentiation between the ignore and play conditions, and meet the criteria for Subtype 1.

Figure 7

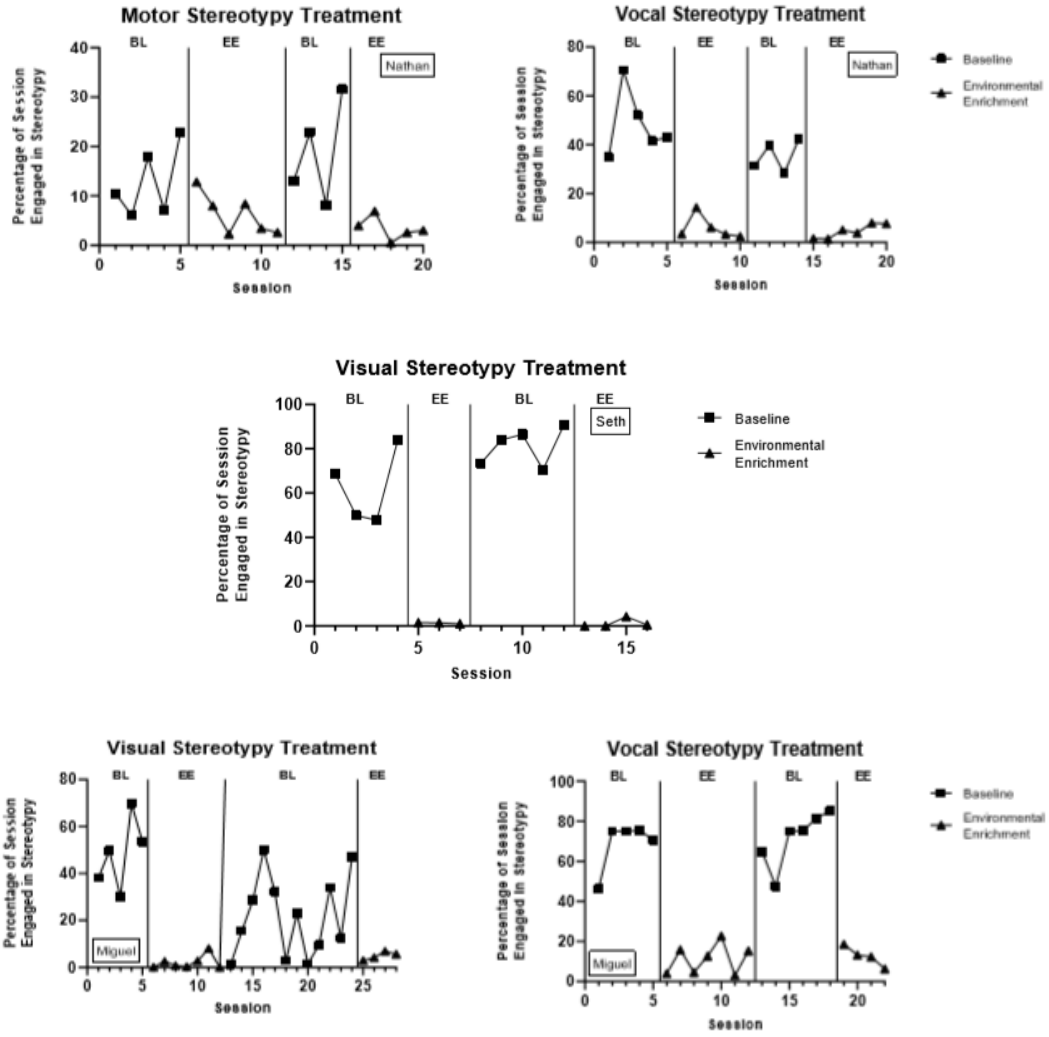

Figure 7. This figure shows treatment data for the three participants' combined total five topographies of stereotypy. All five topographies of stereotypy analyzed met the subtype 1 criteria, and were responsive to environmental enrichment. None of the topographies analyzed required a punishment assessment or punishment as a treatment.