Differences in Immediate and Subsequent Effects of Stimuli on Vocal Stereotypy Inform DRO and DRA Treatments

by

Kristen M. Brogan

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 5, 2017

Keywords: autism, competing stimulus assessment, FOCSA, preference assessment, stereotypy

Approved by

John T. Rapp, (Chair), Professor, Psychology
Sacha T. Pence, Assistant Professor, Psychology
Chris Newland, Professor, Psychology

Abstract

We conducted a series of studies on the immediate and subsequent effects of stimuli on vocal stereotypy exhibited by four children with disabilities. Results of Study 1 showed that vocal and motor stereotypy of all four participants persisted in the absence of sosial consequences. Results of Study 2 showed that a free-operant competing stimulus assessment (FOCSA) identified a high-preference low-stereotypy (HP-LS) stimulus, which was predicted to decrease stereotypy, and a high-preference high-stereotypy (HP-HS) stimulus, which was predicted to increase stetreotypy, for each participant. Results of Study 3 showed that the FOCSA correctly predicted the immediate effect of the HP-LS stimulus for all 4 participants; however, the FOCSA predictions were less accurate for the HP-HS stimulus. Results of Study 4 showed that a differential reinforcement of other behavior procedure in which participants earned access to the HP-LS for omitting vocal stereotypy increased each participants' latency to vocal stereotypy; however, clinically significant durations were only achieved for one participant. Study 5 showed that differential reinforcement of alternative behavior in which participants earned access to the HP-LS stimulus contingent upon correct responses during discrete trial training reduced levels of stereotypy and increased correct engagement for all participants. The potential utility of the FOCSA is discussed.

Acknowledgements

I would like to thank John Rapp for his invaluable guidance and mentorship throughout this project. I would also like to thank my thesis committee members, Sacha Pence and Chris Newland, for their input and assistance. Lastly, I would like to thank Jodi Coon, Anna Kate Edgemon, Amanda Niedfeld, Bailey Sturdivant, Adam Almanza, Alexandra Castre and Aileen Solero for their assistance with data collection throughout this project.

Table of Contents.

List of Tables

List of Figures

List of Abbreviations

ABA	Applied behavior analysis
AO	Abolishing operation
ASD	Autism Spectrum Disorder
DRO	Differential reinforcement of other behavior
DTT	Discrete trial training
DRA	Differential reinforcement of alternative behavior
EO	Establishing operation
FOCSA	Free operant competing stimulus assessment
HP-LS	High preference low stereotypy
HP-HS	High preference high stereotypy
IOA	Interobserver agreement
LP-LS	Low preference low stereotypy
LP-HS	Low preference high stereotypy
MTS	Momentary time sampling
MO	Motivation operation
NCR	Noncontingent reinforcement
NI	No interaction
TCMS	Three component multiple schedule

Introduction

Topographically repetitive or stereotyped behavior is deemed problematic because it may compete with an individual's engagement in socially appropriate interactions and academic behavior (e.g., Enloe & Rapp, 2014; Lanovaz, Roberston, Soerono & Watkins, 2013). Examples of topographically repetitive or stereotyped behavior include body rocking, limb posturing, hand flapping, repetitive grunting and noncontextual humming (Lanovaz & Sladeczek, 2012; Rapp & Vollmer, 2005). When comparing the repetitive behavior of typical children to that of children diagnosed with a developmental disability, researchers have characterized the latter group's stereotypic behavior as developmentally and socially inappropriate (Cunningham & Schreibman, 2008). Furthermore, individuals diagnosed with autism exhibit repetitive behavior that is more overt and frequent than those diagnosed with other developmental disabilities (Cunningham & Schreibman, 2008). If stereotypy interferes with an individual's development of social and academic skills, it may lead to isolation from peers and caregivers (Cunningham & Schreibman, 2008; DiGennaro Reed, Hirst, & Hyman, 2012; Lanovaz & Sladeczek, 2012). On a broader level, children's engagment in high levels of restrictive and repetitive behavior, which include stereotypy, has been associated with increased caregiver stress (Harrop, McBee & Boyd, 2016). Although stereotypy has been researched by various disciplines over the past 30 years, it remains a difficult class of behavior to treat for individuals with ASD.

The context in which an individual engages in stereotypy directly influences the behavioral intervention that a practitioner uses to decrease it (Rapp & Lanovaz, 2016). For example, if treatment takes place during instructional periods, a practitioner may choose an intervention that does not interrupt instruction. Alternatively, if treatment takes place during free time, a practitioner may not be as concerned with whether the intervention interrupts the current activity (Rapp, Cook, McHugh & Mann, 2017). It is important for practitioners to have different treatment options so that they can choose the best treatment for the setting. Three such empirically supported behavioral interventions for stereotypy are DRO (e.g., Lanovaz & Argumedes, 2010; Lanovaz et al., 2014; Rapp et al., 2017), NCR with matched stimulation (e.g., Lanovaz, Fletcher, & Rapp, 2009; Lanovaz & Sladeczek, 2012; Piazza, Adelinis, Hanley, Goh, & Delia, 2000; Rapp, 2006, 2007; Rapp et al., 2013, 2017), and DRA (DiGennaro Reed et al., 2012; Rapp & Vollmer, 2005).

When implementing DRO, a practitioner delivers preferred stimuli to the participant contingent upon the omission of stereotypy for a predetermined time interval (Rapp et al., 2017; Vollmer & Iwata, 1992). For example, a practitioner may provide access to music if the participant is able to refrain from engaging in stereotypy for a continuous duration of 30 s (Lanovaz et al., 2014). Because DRO requires the omission of stereotypy for a continuous period of time to gain access to preferred stimuli, the procedure may lend itself well to treatment of stereotypy during instructional periods.

To implement NCR with matched stimulation, practitioners provide continuous access to stimuli that produce sensory output that resembles the hypothesized sensory output of stereotypy (Piazza et al., 2000). For example, if an individual engages in vocal stereotypy for which the putative reinforcer is auditory stimulation, a therapist may provide that individual with continuous access to music (e.g., Lanovaz et al., 2009; Lanovaz, Sladcezek, & Rapp, 2012; Rapp, 2007). For treatment during free or leisure periods, NCR with matched stimulation may be more appropriate because the stimulation may be disruptive to other activities such as a therapist's delivery of instruction. Because access to matched stimuli is provided continuously without a response requirement, it can be described as NCR (e.g., Richman, Barnard-Brak, Grubb, Bosch, & Abby, 2015).

Researchers have shown the effictiveness of DRA to reduce stereotypy (DiGennaro Reed et al., 2012; Rapp & Vollmer, 2005). The main advantage of DRA over DRO, is that an alternative, more socially appropriate behavior is targeted for increase in a DRA (Lanovaz et al. 2014). Lanovaz et al. (2014) evaluated the effects of NCR with music and DRA on the vocal stereotypy of 12 individuals with developmental disabilities. Overall, results showed decreases in vocal stereotypy from both treatments, with larger decreases from the NCR preparation. However, there are some situtations in which continuous music may not be an option. As such, further investigation of a DRA for the treatment of vocal stereotypy is warranted.

When implementing behavior change procedures, it is important for practitioners to consider current MOs that may influence the effectiveness of an intervention (Cooper, Heron & Heward, 2007). Researchers have utilized the MO concept as a tool to identify preferred stimulation that is functionally matched to stereotypy (e.g., Lanovaz & Argumedes, 2010; Rapp, 2006, 2007; Rapp et al., 2013, 2017). Specifically, researchers have evaluated changes in participants' motivation to engage in stereotypy by measuring stereotypy before, during, and after intervention (Enloe & Rapp, 2014; Frewing, Tanner, Bonner, Baxter & Pastrana, 2015; Lanovaz & Argumedes, 2010; Lanovaz et al., 2009; Lanovaz, Rapp & Fletcher, 2010; Lanovaz & Sladeczek, 2012; Morrison, Roscoe & Atwell, 2011; Pastrana, Rapp & Frewing, 2013; Rapp, 2006, 2007; Rapp et al., 2013, 2017). Briefly, Laraway, Snycerski, Michael, and Poling (2003) defined MOs as stimulus events that alter both the value of a stimulus as a reinforcer or punisher and the frequency or likelihood of behavior that has produced access to that stimulus in the past. When studying changes in motivation in stereotypy, Rapp (2008) suggested that the valuealtering and behavior-altering effects of MOs were inextricably linked. That is, changes in the latter, which were observable, could be assumed to be indicative of changes in the former, which may be less accessible.

As the MO concept pertains to stereotypy, an EO increases the value of sensory stimulation produced by stereotypy as a reinforcer and evokes engagement in stereotypy. Conversely, an AO decreases the value of sensory stimulation produced by stereotypy as a reinforcer and abates engagement in stereotypy (Lanovaz et al., 2010). Although NCR with matched stimulation has shown some promise in the production of possible abolishing effects, the subsequent effects of evocative stimuli remains largely unstudied (Lanovaz, Rapp, Long, Richling, & Carroll, 2014; Rapp et al., 2013). If access to a stimulus increases immediate engagement in stereotypy, these immediate increases may lead to an AO for subsequent engagement in stereotypy. A stimulus that creates a subsequent AO for further engagement in stereotypy may provide practitioners another option for treating stereotypy. When embedded within a broader intervention such as DRO, the subsequent AO effect may allow clinicians and researchers to quickly increase DRO intervals to practical variations (e.g., DRO 5 min). If successful, children and adults who engage in stereotypy that interferes with activities in applied settings may benefit from these findings. Moreover, such findings can potentially extend our conceptual understanding of non-social reinforcement.

The purpose of the current series of studies was to evaluate the effictiveness of NCR with matched stimulation, DRO and DRA treatments using stimuli that had various effects on vocal stereotypy. Specifically, we evaluated the effects of stimuli that either increased stereotypy or decreased stereotypy as identified by a competing stimulus assessment. The targeted form of stereotypy across was vocal stereotypy, however, previous research has shown that decreases in one form of stereotypy may lead to reallocation to other forms of stereotypy (e.g., Rapp, 2005; Rapp et al., 2013; Rapp, Vollmer, St. Peter, Dozier & Cotnoir, 2004). As such, researchers in the current experiment tracked a unique form of nontargeted motor stereotypy for each participant across studies.

In Study 1 we verified that each participant engaged in repetitive vocalizations that persisted in the absence of social consequences. In other words, Study 1 was conducted to verfiy that each participant engaged in vocal stereotypy. In Study 2 we conducted a FOCSA to evaluate the immediate effects of age-appropriate toys on targeted (vocal) and nontargeted stereotypy. In Study 3 we conducted a brief treatment evaluation to evaluate the predictive validity of the FOCSA. Specifically, the treatment evaluation in Study 3 examined both immediate and subsequent effects of two categories of preferred stimuli on targeted and nontargeted stereotypy. In Study 4 we implemented a DRO treatment embedded in DTT sessions using the stimulus that was validated in Study 3 as the reinforcer for meeting DRO intervals. In Study 5, we evaluated the effectiveness of a DRA procedure for decreasing stereotypy in which the participant earned access to the stimulus validated in Study 3 for correct responding on tasks during DTT sessions.

Study 1: Verifying a Non-Social Function

Vollmer, Marcus, Ringdahl and Roane (1995) evaluated a hierarchial functional analysis in which the final step was a series of consecutive no interaction sessions, utilized to identify automatically maintained behavior. Querim et al. (2013) emphasized that automatically maintained behavior generally persists in alone conditions of a functional analysis (FA) whereas behavior that is maintained by social consequences is extinguished in these conditions. Based on these expected patterns, Querim et al. hypothesized that consecutive no-interaction conditions may take the place of a full standard FA (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994) in the assessment of automatically maintained behaviors. Querim et al. conducted a series of 5-min no interaction sessions and found that these sessions accurately predicted the function of 28 out of 30 cases of problem behavior. The sequence evaluated by Querim et al. allows researchers to confirm whether a behavior persists in the absence of social consequences in a brief format. Researchers have adopted this as a suitable procedure to confirm that a behavior is automatically maintained (Enloe & Rapp, 2014; Rapp et al., 2017). As such, Study 1 utilized this procedure to confirm that participants' behavior persisted without social reinforcement.

Method

Participants, setting and response definitions. Four male children participated in the study. Mark was a 3-year-old boy diagnosed with ASD. At the time of study entry, Mark did not have any verbal communication skills or prior training with ABA procedures or instruction. Mark was not receiving any ASD specific services outside of the study. Nick was a 4-year-old diagnosed with ASD who communicated with one to two word mands and had approximately one year of ABA therapy through a University clinic. Nick attended a school for children with disabilities where he received ABA and speech therapy. Sam was a 4-year-old diagnosed with ASD who communicated with a picture exchange system. Sam was able to independently request approximately five different preferred food items with the picture exchange system. Sam had received approximately 1 year of ABA therapy through a University clinic and attended a public preschool special education program where he received ABA and speech therapy. Alex was a 3-year-old diagnosed with a speech delay. Alex did not have any verbal communication skills or prior experience with ABA. Alex attended a public preschool special education program where he received speech therapy.

Sessions were conducted in a 1.5 m x 4.5 m or a 3.6 m x 3.6 m therapy room in a University clinic. The 1.5 m x 4.5 m room contained only a file cabinet and table. The 3.6 m x 3.6 m room was devoid of all materials but included a one way mirror on one wall. Table 1 contains response definitions for each participants' targeted vocal stereotypy and nontargeted motor stereotypy (various response forms).

Procedures. Mark's parents enrolled him in the study when he was beginning clinic services at the University. As such, students in a Master's program for ABA conducted a standard functional analysis (FA; Iwata et al. 1982/1994) for Mark as part of their training. Because data from the standard FA were undifferentiated (data not depicted), researchers subsequently conducted consecutive no interaction (NI) sessions to verify the non-social function (Querim et al., 2013; Vollmer et al., 1995). All of Mark's sessions were 10 min in duration and occurred across three days with five to eight test sessions per day and two to four NI sessions per day. The remaining participants were exposed to consecutive NI sessions only. These consecutive NI sessions were 5 min in duration and researchers conducted two to five sessions per day across two to five days. Mark participated in six NI sessions, Nick participated in five NI sessions, Sam participated in 13 NI sessions and Alex participated in nine NI sessions. Researchers continued NI sessions until vocal stereotypy persisted at stable percentages.

Data collection and reliability. Researchers took data on targeted and nontargeted responses using 10-s MTS (Devine, Rapp, Testa, Henrickson & Schnerch, 2011; Gardenier, MacDonald & Green, 2004; Meany-Daboul, Roscoe, Bourret & Ahearn, 2007; Rapp, Colby-Dirksen, Michalski, Carroll & Lindenberg, 2008; Rapp et al., 2007) either in session or from video recordings of sessions. We converted raw data to a percentage of intervals with stereotypy measure by dividing the number of intervals in which stereotypy was scored by the total number of intervals in the session (60 for Mark or 30 for the remaining participants) and multiplied the result by 100. A secondary observer scored at least 40% of sessions. Researchers then calculated IOA using the interval-by-interval method of agreement for MTS data collection (Mudford, Taylor & Martin, 2009) by checking for agreement in each 10-s interval and dividing the number of intervals with agreement by the total number of intervals. Researchers then multiplied that number by 100. Mean IOA scores for Mark's vocal and motor stereotypy were 98% and 97%, respectively. Mean IOA scores for Nick's vocal stereotypy, stomping, posturing, pacing and facial was 91%, 97%, 98%, 92% and 88%, respectively. Mean IOA scores for Sam's vocal steretoypy, facial, bruxing, pacing, hand posturing, and other motor was 89%, 97%, 99%, 93%, 93%, and 93%, respectively. Mean IOA scores for Alex's vocal stereotypy, hand stereotypy, pacing, and surface rubbing was 93%, 89%, 93% and 92%, respectively.

Results and Discussion

Figure 1 shows that Mark (first panel) engaged in high percentages of both vocal stereotypy (*M* = 85%) and motor stereotypy (*M* = 78%) throughout the consecutive NI sessions. Nick (second panel) engaged in moderate to high percentages of vocal stereotypy (*M* = 56%) and pacing (*M* = 66%), and low percentages of stomping (*M* = 11%), posturing (*M* = 5%) and facial stereotypy (*M* = 12%). Sam (third panel) engaged in moderate percentages of vocal stereotypy (*M* = 42%), moderate to low percentages of pacing (*M* = 32%), hand posturing (*M* = 23%) and surface rubbing (*M* = 24%), and low percentages of bruxing (*M* = 1%), facial stereotypy (*M* = 3%) and other motor (*M* = 7%). Alex (fourth panel) engaged in moderate percentages of pacing (*M* = 43%), moderate to low percentages of vocal stereotypy (*M* = 27%) and surface rubbing (*M* = 37%) and low percentages of hand stereotypy (*M* = 10%).

Results show that each participants' repetitive behaviors persisted in the absence of social consequences, thus meeting the definition of vocal stereotypy (Lanovaz & Sladeczek, 2012). Additionally, each participant engaged in at least one form of repetitive or invariant motor movement that maintained in the absence of social consequences, thus meeting the definition of motor stereotypy (Rapp & Vollmer, 2005).

Study 2: Using a Free-Operant Competing Stimulus Assessment to Identify Stimuli That Decrease and Increase Vocal Stereotypy

Currently, there is no consensus amongst researchers about how to best identify stimuli to be used in the treatment of stereotypy. One option is to conduct a competing stimulus assessment to identify one or more items that compete with stereotypy; however, there is no one agreed upon assessment for which to do so. Frewing, Rapp, and Pastrana (2015) described a FOCSA comprised of three 10-min sessions to identify stimulus preference and conditional percentages of stereotypy when an individual was engaging with a stimulus. A FOCSA yields valuable information pertinent to the treatment of stereotypy, however, only one published study to date has utilized this procedure. As such, more research is needed to validate the FOCSA as an efficient assessment tool to identify stimuli to be used in the treatment of stereotypy. A FOCSA is an adaptation of the free operant stimulus preference assessment (Roane, Vollmer, Ringdahl & Marcus, 1998), but it yields slightly different information.

Although Roane et al. (1998) described collecting data on item engagement and engagement in problem behavior, the FOCSA preparation described by Frewing et al. (2015) prescribes calculting conditional measures of target behavior given specific item engagement.

By calculating these conditional measures of stereotypy, researchers can determine both the individual's relative preference for each item and the level of stereotypy associated with each item. Two of the classifications yielded from the FOCSA are the HP-LS classification and the HP-HS classification. An item identified as HP-LS results in lower levels of stereotypy while the participant is engaging with the item when compared to the background percentage of stereotypy (Frewing et al., 2015). By contrast, an item identified as HP-HS results in higher levels of stereotypy while the participant is engaging with the item when compared to the background percentage of stereotypy (Frewing et al., 2015).

It is not yet known how an HP-HS item may be used in the treatment of stereotypy. Several studies have shown that ambient stimulation can increase individuals' engagement in one or more forms of stereotypy (e.g., Cook, Rapp, Gomes, Frazer, & Lindblad, 2014; Rapp, 2004, 2005; Van Camp et al., 2000); however, no study has attempted to utilize the relation between evocative stimuli as a treatment for stereotypy. The increased levels of stereotypy when tangible stimuli such as toys, or music were made available in the preceding studies suggest that ambient stimulation may act as an EO for engagement in stereotypy in some cases. If ambient stimuli act as an EO for engagement in stereotypy, engagement in stereotypy may decrease below baseline levels when researchers remove the evocative stimulus. Moreover, after such a decrease, it may take more time for stereotypy to return to baseline levels than if the stimulus had not evoked stereotypy. Further research on stimuli that evoke stereotypy is justified by the prevalence of the phenomenon and the limited directed research. The purpose of Study 2 was to identify one HP-LS stimulus and one HP-HS stimulus or a LP-HS stimulus if a HP-HS was not identified.

Method

Participants, setting, and response definitions. All participants, settings and stereotypy response definitions in Study 2 were the same as in Study 1. In addition, researchers recorded data on item engagement, defined as contact between an item and any part of the participant's body for non-auditory items (e.g., plastic figurine) and contact between an item and any part of the participant's body and/or eye-gaze directed towards an item that was producing noise for items that had the ability to make noise (e.g., iPad™).

Data collection and reliability. Researchers used 10-s MTS to record data on targeted stereotypy, nontargeted stereotypy (see Table 1), and item engagement. Each 10-min session contained 60, 10-s intervals. For each 10-s interval, researchers scored item engagement on the 8th second and stereotypy on the 10th second. For example, if a researcher scored item engagement on the 8th second and stereotypy on the 10th second, the interval that comprised seconds 1 through 10 would have been scored for both item engagement and stereotypy. Researchers scored data in this way to ensure accurate measurement of the dependent variables. We converted data to a percentage of intervals with stereotypy by dividing the number of intervals in which stereotypy was scored by the total number of intervals in the session (60) and multiplied the result by 100 to obtain background percentage of sterotypy. We also converted data on item engagement by dividing the number of intervals in which item engagement was scored by the total number of intervals in the session and multiplied the result by 100 to identify item preference. Additionally, we calculated the conditional percentage of targeted stereotypy given item engagement by counting the number of intervals in which both item engagement and stereotypy occurred. We then divided that number by the total number of intervals in which item engagement occurred, regardless of stereotypy and multiplied that number by 100.

A secondary observer collected data for at least 33% of FOCSA sessions across participants. For Mark, mean IOA scores for targeted stereotypy, nontargeted stereotypy, and item engagement were 93%, 96%, and 98%, respectively. For Nick, mean IOA scores for targeted stereotypy, nontargeted stereotypy, and item engagement were 91%, 98%, and, 99% respectively. For Sam, mean IOA scores for targeted stereotypy, nontargeted stereotypy, and item engagement were 95%, 91%, and, 99% respectively. For Alex, mean IOA scores for targeted stereotypy, nontargeted stereotypy, and item engagement were 95%, 98%, and, 99% respectively.

Procedures and design. Researchers conducted three 10-min sessions in which participants had free access to three matched stimuli and three unmatched stimuli. Piazza et al. (2000) described a stimulus as matched when it produced the same overt sensory stimulation that is hypothesized to maintain an automatically reinforced behavior. By contrast, an unmatched stimulus would be a stimulus that does not produce that sensory stimulation (Piazza et al., 2000). As the present study focused on vocal stereotypy and the hypothesized sensory stimulation to maintain vocal stereotypy is auditory stimulation, matched stimuli were those that produced auditory stimulation (e.g., an iPad™ that played children's songs, figurines such as cars or popular characters or animals that produced noise when their buttons were pressed, and a children's piano). Unmatched stimuli used in the present study included a wooden train set, a robot toy, and popular character figurines.

For all participants, we conducted one 10-min session per day, across three days (Frewing et al., 2015). During the initial assessment, Mark, Nick, and Sam simultaneously manipulated the iPad™ and at least one other item. As such, an additional series of three 10-min sessions across three days was conducted in which access to the iPad™ was not provided. The purpose of the iPad™ restriction condition was to identify one item that increased levels of stereotypy beyond baseline percentages. Researchers attempted to identify an item associated with increased levels of stereotypy so that the subsequent effects of this type of item could be examined in Study 3. Specifically, we used the classifications outlined by Frewing et al. (2015). A HP-LS item was in the top 50% of item engagement and associated with a low conditional percentage of stereotypy when compared to both the background percentage of stereotypy (throughout the FOCSA sessions) and the NI sessions from Study 1. A HP-HS item was in the top 50% of item engagement and associated with a high conditional percentage of stereotypy when compared to both the background percentage of stereotypy (throughout the FOCSA) and NI sessions. A low-preference, low-stereotypy (LP-LS) item is one that falls in the bottom 50% for item engagement and is associated with a low conditional percentage of stereotypy. A LP-HS item was in the bottom 50% of item engagement and associated with a high conditional percentage of stereotypy.

Results and Discussion

Figure 2 shows the results of the FOCSA for Mark. During Mark's FOCSA with iPad™ series, the percentage of iPad™ engagement fell within the top 50% of items provided and thus was labeled as a high preference item. Mark also engaged with the ship toy for high percentages of intervals. Mark emitted low to zero percentages of engagement with all other stimuli. Both the iPad™ and the ship were associated with low conditional percentages of stereotypy; however, Mark engaged with both items simultaneously and thus, one item may have been solely responsible for the lower levels of stereotypy. Results of the FOCSA with iPad™ restriction show that the ship fell in the top 50% of items for engagement and that the ship was associated with a high conditional percentages of stereotypy. These results taken together suggest that the ship met the criteria for the HP-HS item. Mark's overall results suggest that the iPad™ was a HP-LS stimulus and that the ship was a HP-HS stimulus (see Table 2).

Figure 3 shows the results of Nick's FOCSA with iPad™. Results show that the iPad™ was associated with high percentages of engagement and low conditional percentage of stereotypy. All other items during this series were associated with low to zero percentages of engagement and low probabilities of stereotypy. Results of Nick's FOCSA with iPad™ Restriction series show that the train fell in the bottom 50% of engagement and was associated with high conditional percentage of stereotypy. Overall results suggest that for Nick, the iPad™ was a HP-LS stimulus and the train was a LP-HS stimulus.

Figure 4 shows the results of Sam's FOCSA with iPad™ indicated that the iPad™ and the train fell in the top 50% of item engagement and were both associated with low conditional percentages of stereotypy. Results of Sam's FOCSA with iPad™ restriction indicated that the train fell in the top 50% of item engagement and was associated with a high conditional percentage of stereotypy. Overall results of Sam's two FOCSA series suggested that the iPad™ was a HP-LS stimulus and the train was a HP-HS stimulus.

The results of Alex's FOCSA with iPad™ (the only one conducted, Figure 5) indicated that the iPad™ fell in the top 50% of item engagement and was associated with a low conditional percentage of stereotypy indicating that it was a HP-LS stimulus. The Octonauts™ figurine, and Buzz Lightyear fell in the top 50% of item engagement and were associated with a high conditional percentage of stereotypy, making each one a HP-HS stimulus. For the purposes of Study 3, we chose to assess the HP-HS item that was associated with the highest percentage of engagement (i.e., Octonauts™).

Table 2 summarizes the FOCSA results. The FOCSA identified an HP-LS stimulus for all participants. For Mark, Sam, and Alex, a HP-HS stimulus was identified but for Nick a LP-HS item was identified instead. Study 3 was designed to evaluate the predictive validity of the FOCSA on immediate effects of HP-LS and HP-HS or LP-HS stimuli on stereotypy. As previously noted, a participant should only manipulate one item at a time during a FOCSA. However, Mark, Nick and Sam engaged in simultaneous manipulation of items that may have confounded results. However, researchers accomplished clear results by conducting a FOCSA with iPad™ restriction. The Frewing et al. (2015) study evaluated the immediate effects of stimuli on stereotypy by way of the FOCSA, but they did not evaluate subsequent effects of these stimuli on stereotypy. Subsequent effects on stereotypy are important because they may dictate how an item is used in treatment. For example, if an item reduces immediate engagement in stereotypy but increases subsequent engagement, this treatment may not be useful before periods of time in which stereotypy is undesirable. Subsequent effects of stimuli identified by the FOCSA has not been evaluated in the literature and as such, directed research is warranted. If the immediate effect of a stimulus was correctly predicted by the FOCSA, researchers evaluated that stimulus in Study 3 which was designed to evaluate immediate and subsequent effects of the stimulus on stereotypy.

Study 3: Validating the FOCSA and Evaluating Subsequent Effects on Targeted and Nontargeted Stereotypy

Although NCR with matched stimulation has shown some promise in the production of possible abolishing effects (e.g., Rapp, 2007), the subsequent effects of evocative stimuli on engagement in stereotypy remains largely unstudied (Lanovaz et al., 2014; Rapp et al., 2013). If access to an HP-HS or LP-HS increases immediate engagement in stereotypy and creates an AO for subsequent engagement in stereotypy, this may provide practitioners another option for treating stereotypy. If an HP-HS or LP-HS created a subsequent AO for stereotypy, this would provide a treatment option for individuals for whom practitioners cannot find a functionally matched stimulus.

For each participant, we evaluated the immediate and subsequent effects of the HP-LS and HP-HS (LP-HS for Nick) during NCR. We opted to use NCR during this assessment because it has been previously shown to produce subsequent changes in motivation as determined by a three component multiple schedule arrangment (e.g., Rapp, 2006, 2007; Rapp et al, 2013, 2017). If the FOCSA results are validated, we would expect to see the data path for the second component of the HP-LS sequence to be differentiated from and below the data path for the control sequence and the converse for the data path of the HP-HS sequence. That is, the data path for the second component of the HP-HS sequence should be differentiated above the data path for the control sequence. Further, we hypothesize that some items may produce subsequent decreases in stereotypy once the item is removed. If an item produces a subsequent decrease in stereotypy, then the data path for the third component of that test sequence will be differentiated below the data path for the third component of the control sequence (Lanovaz et al., 2010).

Method

Participants, setting, and response definitions. All participants, settings and stereotypy response definitions in Study 3 were the same as in Study 2. In addition, researchers recorded data on item engagement (same definition from Study 2), when applicable.

Data collection and reliability. Researchers collected data on both targeted and nontargeted stereotypy and item engagement as described in Study 2. A secondary data collector collected data for at least 33% of sessions across participants. Mean IOA scores for Mark’s targeted stereotypy, nontargeted stereotypy and item engagement were 92%, 87% and 100%, respectively. Mean IOA scores for Nick's targeted stereotypy, nontargeted stereotypy and item engagement were 92%, 94% and 99%, respectively. Mean IOA scores for Sam's targeted stereotypy, nontargeted stereotypy and item engagement was 98%, 99% and 98%, respectively. Mean IOA scores for Alex's targeted stereotypy, nontargeted stereotypy and item engagement was 93%, 95% and 97%, respectively.

Procedures and design. Study 3 employed a three-component multiple schedule (TMCS; Lanovaz et al., 2010) to examine the immediate and subsequent effects of noncontingent access for HP-LS and HP-HS or LP-LS stimuli on participants' stereotypy. A TCMS is a variation of a multielement design in which researchers conduct one sequence per day which consists of a series of three, 10-min components. During the TCMS, researchers collected data on stereotypy (both targeted and nontargeted) and item engagement (when applicable). The TCMS employed in the current study involved three sequences; one control sequence and two test sequences. The control sequence was comprised of three, 10-min components of NI. The control sequence allowed researchers to evaluate trends in stereotypy engagement over time when an individual had free access to engage in stereotypy. During the test sequences, components one and three were 10-min NI conditions and component two was 10 min of intervention. The first, second, and third components of the control sequence are then compared to the first, second, and third components of the test sequences, respectively. All participants were exposed to a HP-LS test sequence and either a HP-HS test sequence (Mark, Sam and Alex) or a LP-HS test sequence (Nick). During component two of a test sequence, noncontingent access to the specified stimulus was provided. Component two allowed researchers to evaluate immediate effects of the test item on stereotypy engagement. The immediate effects allow researchers to determine whether or not the FOCSA made accurate predictions of the test item's effects on stereotypy. Component three of a test sequence allowed researchers to evaluate subsequent effects of a test item on stereotypy. In other words, the third component showed levels of stereotypy after access to the test item was discontinued. These subsequent levels of stereotypy reflect potential MOs that the stimulus from the previous component may exert on stereotypy after the stimulus has been removed (Enloe & Rapp, 2014; Lanovaz et al., 2010; Rapp, 2006, 2007; Rapp et al., 2013).

We used visual analysis to determine the immediate and subsequent effects of stimuli throughout the TCMS. Specifically, we focused on the second component wherein we compared each test sequence to the control sequence. We conducted at least three sessions of each sequence. Similar to Rapp et al. (2017), we continued conducting sequences of the TCMS until there was clear differentiation of targeted stereotypy for at least three sessions or a maximum of five sessions were conducted with each sequence (this termination criterion was never applied).

Results and Discussion

Figure 6 shows the results on targeted stereotypy (top panel) and nontargeted stereotypy (middle panel) and item engagement (bottom panel) for Mark. In the targeted stereotypy graphs (top panel) the data paths in component one show undifferentiated levels of stereotypy across all conditions. This is as expected as component one was always a NI condition. Mark's control data path in component two shows relatively high and somewhat variable levels of sterotypy. Mark's HP-LS sequence data path in component two shows low, stable percentages of stereotypy and differentiatiation below the control sequence. Mark's LP-HS sequence data path shows moderate, stable percentages of stereotypy. The data paths in component three show undifferentiated levels of stereotypy across all conditions. These results taken together suggest that the FOCSA accurately predicted the iPad™ as a HP-LS stimulus, but that it did not exert any abative effects after noncontingent access was removed.

The middle panel of Figure 6 shows percentages of motor stereotypy in the HP-LS test condition remained at differentially low levels for Mark in component two. No subsequent effects of either test item were seen for motor stereotypy in Mark. The bottom panel of Figure 6 shows that the FOCSA accurately predicted high levels of engagement with both the HP-LS stimulus and the HP-HS stimulus. Further, the results of the FOCSA showed slightly higher percentages of item engagement with the HP-LS stimulus than the HP-HS stimulus, which was validated in the TCMS.

Figure 7 shows the results on targeted stereotypy (top panel) and nontargeted stereotypy (middle panel) and item engagement (bottom panel) for Nick. In the targeted stereotypy graphs (top panel), the data paths in component one were undifferentiated across all conditions for targeted stereotypy. Nick's control sequence data path in component two shows moderate, stable percentages of targeted stereotypy. Nick's HP-LS sequence data path in component two shows differentially lower, stable percentages of targeted stereotypy suggesting that the HP-LS reliably reduced Nick's immediate engagement in stereotypy when compared to no interaction conditions. Nick's LP-HS sequence data path shows moderate to high variable levels of targeted stereotypy suggesting that the LP-HS may increase levels of immediate stereotypy when compared to no interaction conditions. Data paths from Nick's third component is relatively undifferentiated across conditions indicating that neither the HP-LS or the LP-HS exerted any strong abative or evocative effects on targeted stereotypy following noncontingent access.

Nick showed low percentages of motor stereotypy (middle panel) in component two of both the control and HP-LS sequences. No subsequent effects of either test item were seen for motor stereotypy in Nick. Additionally, item engagement data (bottom panel) show that the FOCSA accurately predicted preference when comparing the HP-LS (consistenly higher percentages of engagement) and LP-HS (consistently lower percentages of engagement).

Figure 8 shows the results on targeted stereotypy (top panel) and nontargeted stereotypy (middle panel) and item engagement (bottom panel) for Sam. Targeted stereotypy in the top panel shows the data paths in component one were undifferentiated across all conditions. Sam's control sequence data path in component two shows moderate, stable percentages of targeted stereotypy. Sam's HP-LS sequence data path in component two shows low, stable percentages of targeted stereotypy that are differentiated below the control. These data suggest that the HP-LS reliably decreased immediate engagement in targeted stereotypy when noncontingent access was provided. The data path in component two of Sam's HP-HS condition shows moderate, stable, undifferentiated percentages of targeted stereotypy. These data suggest that noncontingent access to the HP-HS did not exert any reliable evocative effects on targeted stereotypy engagement. Data paths in component three of Sam's control condition show moderate, stable percentages of targeted stereotypy. Data paths in component three of Sam's HP-LS sequence show differentiated low percentages of targeted stereotypy suggesting that noncontingent access to the HP-LS exerted abative effects on subsequent engagement in vocal stereotypy following noncontingent access. Data paths in component three of Sam's HP-HS sequence show moderate, stable and undifferentiated percentages of targeted stereotypy suggesting that the HP-HS did not exert any effects on subsequent engagement in stereotypy. Percentages of nontargeted stereotypy (middle panel) in the HP-LS test condition remained at differentially low levels during component two for Sam. No subsequent effects of either test item were seen for nontargeted stereotypy in Sam.

The bottom panel of Figure 8 shows Sam's percentage of item engagement during the second component throughout the TCMS. Results show high percentages of item engagement for both HP-LS and HP-HS stimuli across sessions. Additionally, item engagement was consistently slightly higher during the HP-HS sequence which indicates that for some individuals, stimuli that evoke stereotypy, may be of higher preference than stimuli that decrease stereotypy. These results suggest that the FOCSA accurately predicted preference in that higher percentages of item engagement were seen for the HP-HS stimulus than the HP-LS stimulus.

Figure 9 shows the results on targeted stereotypy (top panel), nontargeted stereotypy (middle panel) and item engagement during the second component (bottom panel) for Alex. In the targeted stereotypy graphs (top panel), data paths for stereotypy in component one were undifferentiated across sequences. The targeted stereotypy data path in component two shows the HP-LS sequence differentiated below the control. The HP-HS sequence is undifferentiated from the control sequence. The targeted stereotypy data path in component three is undifferentiated across sequences.

In Alex's nontargeted stereotypy graphs (Figure 9, middle panel), the nontargeted stereotypy data in component one is undifferentiated across sequences. The nontargeted stereotypy data in component two of the HP-LS sequence is differentiated below the control. The HP-HS sequence is undifferentiated. The nontargeted stereotypy data in component three is undifferentated across sequences. The bottom panel of Figure 10 shows the percentage of item engagement for Alex throughout the TCMS. Results show that the HP-LS item was associated with high percentages of item engagement throughout the TCMS. The HP-HS stimulus was associated with high percentages of item engagement throughout the first three HP-HS sequences and a descending trend in engagement is seen in the last few sequences. Overall, the results suggest that the FOCSA accurately predicted preference for Alex.

Taken together, results of Study 3 show several strong points of the FOCSA. First, the FOCSA accurately predicted the immediate effects of HP-LS stimuli for all participants. In other words, the current data set suggests that the predictive validity of the FOCSA for the HP-LS item is strong and as such, the FOCSA may be a useful assessment tool for clinicians who aim to decrease stereotypy. Second, although the HP-LS item was not structurally matched to any of our participants' nontargeted stereotypy, lower percentages of motor stereotypy were seen in the second component for all participants. As such, the HP-LS item may have added desirable effects on nontargeted stereotypy for some individuals. Third, the FOCSA accurately predicted relative preference of each participants' identified stimuli. The FOCSA showed that Mark, Nick and Alex engaged with the HP-LS item for higher percentages than the HP-HS (LP-HS for Nick). Item engagement data from the TCMS show that this held up over time. Additionally, the FOCSA showed that Sam engaged with the HP-HS stimulus for higher percentages than the HP-LS stimulus. This effect held up throughout Study 3. Although the FOCSA may be a valuable assessment tool used to identify an HP-LS item, the FOCSA did not accurately predict the effects of the HP-HS stimuli for any participant.

Subsequent effects of preferred items on stereotypy may be just as important as immediate effects in certain situations. Although noncontingent access to matched stimulation has been shown to be an effective treatment for stereotypy (Lanovaz & Sladeczek, 2012; Piazza et al., 2000; Rapp, 2006, 2007), subsequent effects may dictate the type of treatment a practitioner employs in certain classroom or home environments. If a stimulus exerts subsequent abative effects on stereotypy, a differential reinforcement of other behavior (DRO) treatment in which practitioners deliver the stimulus with continued abative properties contingent on meeting the DRO criterion may be appropriate. However, if that stimulus does not have continued abative effects, this DRO arrangement may not be appropriate.

Study 4: Differential Reinforcement of Other Behavior (DRO) with access to the High-Preference, Low-Stereotypy Stimulus as the Reinforcer

One differential reinforcement procedure that may be used to decrease undesirable behavior is a DRO. During a DRO, a practitioner provides access to a reinforcer contingent upon the omission of the target behavior for a predetermined duration (Vollmer & Iwata, 1992). Researchers have used DRO to decrease stereotypy, however some procedural variations still exist. Gehrman, Wilder, Forton, and Albert (2017) described a DRO 10-s preparation in which an arbitrary reinforcer was delivered for the omission of hand-flapping, which was evoked by ambient stimulation (i.e., YouTube™ videos on a computer). Although this procedure effectively decreased the percentage of intervals with hand-flapping, several limitations should be noted. First, the reinforcer used by Gehrman et al. (2017) was arbitrary. Recent developments in the literature allow researchers and practitioners to provide functional reinforcers for stereotypy in the form of functionally matched stimulation. Second, the 10-s DRO interval was never thinned which may lead to difficulties in clinical implementation (e.g., see Rapp & Lanovaz, 2016 for a review). Rapp et al. (2017) described a trial-based DRO in which a functionally matched stimulus identified by a FOCSA was delivered contingent upon omitting stereotypy for a pre-determined duration. Rapp et al. thinned the DRO interval up to 10 min for one participant.

The purpose of Study 4 was to examine the efficacy of a DRO arrangement in which the HP-LS stimulus was delivered to the participant contingent upon meeting the DRO criterion.

We did not evaluate the HP-HS or LP-HS stimuli because none produced either an immediate increase in vocal stereotypy or a subsequent decrease in vocal stereotypy. We originally hypothesized that in order to produce meaningful treatment effects in a DRO preparation, that a HP-HS stimulus should produce immediate increases and subsequent decreases in stereotypy engagement. Additionally, we predicted that a functionally matched item that produced both an immediate and subsequent decreases would lead to greater treatment gains in a DRO preparation (Rapp et al., 2017). In relation to the current study, Sam's TCMS data showed that the structurally matched HP-LS produced an immediate and subsequent decrease. Given the immediate and subsequent decreases, the HP-LS is a functionally matched stimulus for Sam's vocal stereotypy. As such, we predicted that Sam would have the most success with treatment in DRO.

Method

Participants, setting, and response definitions. Mark, Nick and Sam participated in Study 4. All settings and stereotypy response definitions in Study 4 were the same as in Study 3.

Data collection and reliability. Researchers collected data on the latency to vocal stereotypy for all participants. A secondary observer collected data for at least 33% of sessions across participants. Researchers calculated IOA using a time-window analysis method where an agreement was scored if the latency recorded by the second observer was within 2 s of the latency recorded by the first observer (Mudford et al., 2009; Rapp et al., 2017). For each session in which latency IOA was collected, researchers divided the number of agreements by the number of agreements plus disagreements and multiplied by 100. Mean IOA scores for latency to vocal stereotypy for Mark, Nick and Sam was 98%, 99%, and 100%, respectively.

Procedures and design. Researchers implemented 10-min sessions with a trial based changing criterion design as described by Rapp et al. (2017). Researchers of the present study delivered the HP-LS stimulus contingent upon the participant meeting the current DRO criterion. Participants received the HP-LS stimulus for 20 s, 30 s or 1 min depending on the current DRO criterion. For criteria less than 30 s, the participant received the HP-LS for 20 s, for criteria less than 1 min, participants received the HP-LS for 30 s and for criteria over 1 min participants received the HP-LS for 1 min.

Researchers set initial criteria based off each participant's mean latency to stereotypy in baseline phases (Rapp et al., 2017). Researchers conducted NI baseline phases and DTT baseline phases. The DTT baseline phases for Study 4 were synonymous with DRO sessions except that no differentiatial consequences were provided for stereotypy engagement. The first criteria was set at half of the last 10 vocal stereotypy inter-response times (IRT), excluding outliers. Successive criteria were modest increases to ensure the participants continued to meet the target. Successive increases were systematically varied in that some occurred in large steps and some were more moderate steps in order to strengthen the experimental control of the changing criterion design (Kazdin, 2011). Criterion to increase a DRO duration was set at 10 consecutive successful trials. Criterion to decrease a DRO duration was set at 10 consecutive failed trials. Modifications were made to criterion for Mark and Nick when progress stagnated around 30 s. Researchers modified criteria to increase a DRO duration to four out of five successful trials as opposed to 10 consecutive trials. Sam began Study 4 after Mark and Nick had experienced more success with the modified criteria for increase and decrease. As such, Sam's criteria for increase and decrease was always four out of five successes or failures, respectively.

Similar to Rapp et al. (2017), researchers delivered age- and developmentally-appropriate instructions throughout DRO durations. Before beginning a trial, the researcher stated, “You can earn the iPad™ for working quietly”. If a participant met the DRO criterion, researchers provided brief access as described above to the HP-LS stimulus and enthusiastic praise for working quietly. If a participant did not meet the DRO criterion, the researcher reset the timer, restated the contingency and began a new trial. Researchers attempted to thin each participants’ DRO schedule to at least 300 s for an increased likelihood of acceptability in clinical settings.

Results and Discussion

Mark’s results are depicted in Figure 10. Mark’s mean IRT in DTT baseline was 6 s, thus his first criterion was set at 3 s. Mark met the first criterion in 10 trials. Mark’s second criterion was set at 6 s and Mark met this criterion in 10 trials. Mark’s third criterion was set at 12 s and Mark met this in 26 trials. Mark’s fourth criterion was set at 24 s and Mark met criterion to decrease in 37 trials. Researchers decreased the criterion back to 12 s and Mark met this in 66 trials. The next criterion was a more moderate increase to 18 s instead of doubling the previous criterion in attempts to increase the likelihood of success. Mark met the 18 s criterion in 38 trials. The next criterion was set at 24 s and Mark met this criterion in 207 trials with an increase to 60 s access to the HP-LS after 93 of those trials. The next criterion was set at 28 s and Mark participated in 167 trials at this criterion. At this point, researchers changed the criterion to increase to four out of five successful trials in an attempt to increase treatment effectiveness. Once the criterion to increase was adjusted, Mark met the criterion in 12 trials. The next criterion was set at 50 s and Mark participated in 10 trials before meeting criteria to decrease. The decreased criterion was set at 28 s and Mark met this criterion in 28 trials. The next criterion was a more gradual increase at 40 s and Mark met termination criterion after five trials and the criterion was decreased back to 28 s. Mark met termination criteria after 12 trials and the criterion was reduced to 24 s.

At this point, Mark had participated in a total of 627 trials of the DRO with HP-LS access that occurred over 70 sessions. This was a total of 700 min across 24 days. The clinical acceptability of the procedure began to wane as a criterion of only 24 s had been obtained across this time. This duration was only sufficient to deliver approximately two discrete trial instructions. As such, this phase of the experiment was discontinued for Mark and he began Study 5.

Results for Nick's DRO are shown in Figure 11. Nick's mean IRT in DTT baseline was 24 s so his first criterion was set at 12 s which Nick met in 10 trials. The second criterion was doubled to 24 s which Nick met in 10 trials. The third criterion was doubled as well to 48 s and Nick met criterion to decrease after 80 trials. The next criterion was set back to 24 s which Nick met in 122 trials. The next criterion was a more moderate increase to 30 s. After 17 trials without meeting criterion to increase, we modified the criterion to increase to four out of five consecutive trials. After this modification, Nick met criteria to increase from 30 s after 19 trials. The next criterion was set at 50 s and Nick met criteria to increase after 16 trials. The next criterion was set at 65 s and Nick met criteria to decrease in 5 trials. The criterion returned to 50 s and Nick met criteria to decrease again after 10 trials. The criterion returned to 30 s which Nick met in 7 trials. The next increase was a more moderate one to 40 s which Nick failed to master and met criteria for decrease after six trials. The criterion was then set back to 30 s which Nick failed to master and met criteria for decrease in five trials. The next criterion was decreased to 25 s and Nick had three successful trials and one failed trial before the session ended.

At this point, Nick had participated in 393 trials over 47 sessions (totaling 470 min of session time) and 21 days without consistent mastery of a duration over 30 s. As with Mark, it seemed unrealistic for practitioners to spend over 7 hr attempting to increase a DRO without yet reaching clinically acceptable levels. Although a trial-based DRO allows for less failures than a session-based DRO because increase and decrease criteria were based on successful and unsuccessful trials as opposed to sessions, the trial-based DRO appeared to lack clinical utility for Nick. At this point, the DRO was discontinued for Nick and he entered into Study 5.

Sam's results are shown in Figure 12. Sam's DTT baseline IRT was 16 s. Because the results of Study 3 showed that the HP-LS exerted abative effects for subsequent engagement in stereotypy, Sam's first criterion was set slightly above his mean baseline IRT at 20 s. Sam met this criterion in 11 trials and the next criterion was set at 30 s. Sam met the 30-s criterion in four trials, the minimum number of trials to master, and the next criterion was set at 45 s. Sam met the 45 s criterion in five trials and the next criterion was set at 90 s. Sam met criterion to decrease in seven trials and the criterion was decreased back to 45 s. Sam met criterion to decrease in five trials and the criterion was decreased back to 30 s. Sam met the 30 s criterion in four trials and the next criterion was set at 60 s. Sam met criterion to decrease in seven trials and the criterion was set back to 30 s which Sam met in seven trials. The next increase was more moderate, to 47 s. Sam met criterion to decrease in seven trials and the criterion was set back to 30 s. Sam met this criterion in four trials and the next criterion was an even more moderate increase to 35 s. Sam met this in five trials and the next criterion was set at 40 s. Sam met criteria for decrease in four trials so the criterion was returned to 30 s. Sam met criteria for decrease in four trials.

At this point, researchers noted that Sam was engaging in vocal stereotypy when he placed the edible he had earned for correct responding during discrete trial tasks in his mouth. The vocal stereotypy was low-magnitude and sounded like a continued emission of the sound "mmm" with a closed mouth. As this was only occurring when Sam placed the edible in his mouth, it was not interfering with teaching and appeared to be evoked by consuming the edible, researchers placed a 10 s hold on consequating stereotypy after edible delivery. In other words, if Sam engaged in the "mmm" behavior, researchers would not re-set the interval if it was within 10 s of receiving an edible for correct responding.

The first criterion after the 10 s hold was implemented was set at 25 s. Sam met this criterion in five trials and the next criterion was set at 40 s which Sam met in four trials. The next criterion was set at 70 s which Sam met in four trials. The next criterion was set at 120 s and Sam met this in four trials. The next criterion was set at 270 s and Sam met this in four trials. The next criterion was set at 300 s and Sam met this in seven trials. With this modification, Sam met a clinically acceptable DRO duration (5 min) in a total of 29 trials which occurred over eight sessions (equaling 80 min total) and three days. When Sam mastered the 300 s criterion, researchers implemented a return to DTT baseline. Sam's highest IRT during this phase was 851 s. We re-introduced the DRO with the HP-LS at 300 s and Sam mastered this criterion in four trials. The next criterion was set at 420 s and Sam mastered this in 11 trials. The next criterion was set at 600 s and Sam met criterion to decrease in five trials. We reset the criterion at 420 s and Sam met criterion for increase in six trials. The next criterion was set at 460 s and Sam met criterion for increase in four trials. The next criterion was set at 506 s and

Sam met criterion for decrease in five trials. The next criterion was set at 460 s and Sam met criterion for decrease in five trials. The next criterion was set at 300 s and Sam met criterion for decrease in four trials. At this point, Sam had participated in 177 trials across 48 sessions (480 min) and 17 days. As such, Sam was moved into Study 5.

The results of the TCMS in Study 3 predicted that Sam would be the most successful with the DRO procedure in that he was the only participant who showed immediate and subsequent decreases in stereotypy after NCR with the HP-LS stimulus. The current data set is consistent with the literature in that the combination of immediate and subsequent abative effects of stimuli on stereotypy may be rare. As such, it seems important that clinicians have a treatment option for individuals who only show immediate effects of the HP-LS. One differential reinforcement procedure that directly teaches other skills that may not require a subsequent decerase is a DRA.

Study 5: Differential Reinforcement of Alternative Behavior with Contingent Access to the HP-LS Stimulus

Caregivers and clinicians may prefer to utilize a DRA treatment for problem behavior as it involves the selection and teaching of an alternative, socially appropriate behavior. Lanovaz et al. (2013) conducted a systematic review in which they discussed the effects of reducing stereotypy on other behaviors. Lanovaz et al. discussed five studies that utilized some form of differential reinforcement to decrease stereotypic behavior and increase alternative socially appropriate behaviors. Some alternative socially appropriate behaviors included the use of an activity schedule, object-contact responses with a micro-switch, and play skills. Lanovaz et al. noted that only one study in their review had targeted correct responding on academic tasks and that study had used a DRO preparation. The lack of the use of a DRA to decrease stereotypy and increase academic skills is surprising as possible detrimental effects of stereotypy on skill acquisition has been noted as a reason to treat stereotypy (Cunningham & Schreibman, 2008; Koegel, Firestone, Kramme, & Dunlap, 1974). As such, the current study took advantage of this lack of directed research and evaluated a DRA to decrease vocal stereotypy in which participants earned the HP-LS stimulus contingent upon correct responses on acquisition tasks during DTT sessions.

Method

Participants, setting, and response definitions. All participants, settings and stereotypy response definitions in Study 5 were the same as in Study 3. As secondary measures, we tracked responding on gross motor imitation tasks to ensure target acquisition and frequency data on demand presentation which was defined as any instance in which the therapist delivered a directive to the student.

Data collection and reliability. Researchers collected data on both targeted and nontargeted stereotypy as described in Study 2. Researchers calculated IOA for targeted stereotypy for at least 78% of sessions across participants. Mean IOA scores for targeted stereotypy for Alex, Nick, Mark and Sam were 93%, 92%, 90%, and 97%, respectively. Mean IOA scores for nontargeted stereotypy for Alex, Nick, Mark and Sam were 99%, 99%, 90%, and 98%. Researchers calcultated IOA for rate of demands for at least 29% of sessions across participants. Mean IOA scores for rate of demand for Alex, Nick, Mark and Sam were 97%, 99%, 96%, and 97%, respectively.

Procedures and design. Researchers used a 4-tiered nonconcurrent multiple baseline across participants design with an ABACACD reversal in the first tier and an ABAB reversal in the third tier to evaluate the effects of DRA with the HP-LS item. Researchers implemented 3-

min sessions throughout Study 5 for Mark, Nick and Sam. Researchers chose 3-min sessions as this duration fit the natural stream of instruction for these participants. Alex participated in 5-min sessions. Baseline conditions were identical to the no-interaction conditions previously described. During DRA with HP-LS, researchers administered gross-motor imitation teaching in the form of DTT. Contingent upon correct responses for acquisition targets, the therapist delivered behavior specific praise and 30 s of access to the HP-LS item (children's music videos on the iPad™ for all participants). During the DRA with edibles condition (Alex only), researchers administered gross-motor imitation teaching in the form of DTT. Contingent upon correct responses for acquisition targets, the therapist delivered behavior specific praise and one small edible (e.g., half of a skittle or half of an M&M, etc.). A DRA with edibles may be another treatment option; however, it is unlikely to be functionally matched to stereotypy. During the DRA with praise condition (Alex only), researchers administered gross-motor imitation teaching in the form of DTT. Contingent upon correct responses, the therapist delivered behavior-specific praise only. A DRA procedure with contingent praise is another possible treatment option that may be functionally matched to vocal stereotypy (Enloe & Rapp, 2014).

Results and Discussion

Results of Study 5 on targeted stereotypy are depicted in Figure 13. Across three baseline sessions, Alex (top panel) engaged in variable and moderate percentages of vocal stereotypy (*M* = 33%). The first treatment phase for Alex was DRA with edibles. During this phase, Alex engaged in low percentages of vocal stereotypy (*M* = 7%). Researchers then reversed back to a baseline phase in which Alex engaged in moderate percentages of vocal stereotypy (*M* = 24%). Researchers then implmented DRA with the HP-LS stimulus. During this phase, Alex engaged in low percentages of vocal stereotypy (*M* = 6%). Researchers then reversed back to a baseline phase in which Alex engaged in moderate percentages of vocal stereotypy (*M* = 41%). Researchers then re-implemented the DRA with HP-LS an Alex engaged in low percentages of vocal stereotypy (*M* = 9%). Finally, researchers evaluated DRA with praise and Alex engaged in moderate, increasing percentages of vocal stereotypy (*M* = 48%). Across five baseline sessions, Nick (second panel) engaged in moderate to high percentages of vocal stereotypy (*M* = 64%). When treatment was implemented, Nick engaged in variable, low to moderate percentages of vocal stereotypy (*M* = 28%). Across seven baseline sessions, Mark (third panel) engaged in moderate percentages of vocal stereotypy (*M* = 44%). When treatment was introduced, Mark engaged in low percentages of stereotypy (*M* = 9%). A return to baseline resulted in a quick increase to baseline percentages of stereotypy (*M* = 60%). The re-introduction of treatment resulted in a quick return to low percentages of stereotypy (*M* = 2%). Across nine baseline sessions, Sam (bottom panel) engaged in moderate percentages of vocal stereotypy (*M* = 51%). When treatment was introduced, Sam's stereotypy quickly decerased to low levels (*M* = 3%). Mean percentages of correct responses across DRA sessions for Alex, Nick, Mark and Sam were 90%, 87%, 89% and 99%, respectively. Mean rate of demand across DRA sessions for Alex, Nick, Mark and Sam were 1.4, 1.4, 1.3 and 3.0, respectively.

Results of Study 5 on nontargeted stereotypy are depicted in Figure 14. Across three baseline sessions, Alex (top panel) engaged in low but increasing percentages of nontargeted stereotypy (*M* = 17%). In the DRA with edibles phase, Alex engaged in zero instances of nontargeted stereotpy. In the first return to baseline, Alex engaged in low percentages of nontargeted stereotypy (*M* = 8%). In the first DRA with HP-LS phase, Alex engaged in zero instances of nontargeted stereotypy. In the second return to baseline, Alex engaged in low percentages of nontargeted stereotypy (*M* = 10%). In the second DRA with HP-LS and the DRA with praise phase, Alex engaged in zero instances of nontargeted stereotypy. Across six baseline sessions, Nick (second panel) engaged in moderate and variable percentages of nontargted stereotypy (*M* = 29%). During the DRA with HP-LS phase, Nick engaged in zero instances of nontargeted stereotypy. Across seven baseline sessions, Mark (third panel) engaged in moderate percentages of nontargeted stereotypy (*M* = 48%). During the first DRA with HP-LS phase, Mark engaged in low percentages of nontargeted stereotypy (*M* = 2%). During the return to baseline, Mark engaged in moderate, increasing percentages of nontargeted stereotypy (*M* = 24%). During the second DRA with HP-LS phase, Mark engaged in low percentages of nontargeted stereotypy (*M* = 1%). During eight baseline sessions, Sam (bottom panel) engaged in moderate, decreasing percentages of nontargeted stereotypy (*M* = 24%). During the DRA with HP-LS phase, Sam engaged in zero instances of nontargeted stereotypy.

Results of Study 5 suggest that the HP-LS decreased both targeted and nontargeted stereotypy and increased desirable behavior for all participants. This procedure may be a viable option to treat vocal stereotypy for individuals who show only immediate decreases or both immediate and subsequent decreases with the HP-LS stimulus. Although all participants' stereotypy decreased during DRA with HP-LS phases relative to baseline, Sam's decrease was the largest. The DRO was also most effective for Sam, suggesting the an HP-LS stimulus that produces immediate and subsequent decreases on stereotypy may lead to many different effective treatment options. Although DRA with edibles decreased for Alex, DRA with praise only did not. This suggests that even though a stimulus may be structurally matched to stereotypy, it may not effectively decrease stereotypy. Further, because a TCMS with DRA with edibles as the test sequence was not conducted, it is unclear as to what kinds of subsequent effects DRA with edibles may have on stereotypy. As such, it is possible that DRA with edibles leads to a subsequent increase in vocal stereotypy whereas access to the HP-LS did not lead to subsequent increases in vocal stereotypy for Alex.

General Discussion

Study 1 showed that each participant's repetitive vocalizations and motor movements persisted in the absence of social consequences. Study 2 showed that the FOCSA identified (a) a HP-LS stimulus for each participant and (b) a HP-HS stimulus for three participants and a LP-HS stimulus for one participant. Study 3 showed that the FOCSA accurately predicted the immediate effects of the HP-LS stimulus on targeted vocal stereotypy for all four participants; however, the FOCSA did not consistently predict the effects of the HP-HS and LP-HS stimuli. Results from Study 3 also showed that the HP-LS decreased nontargeted motor stereotypy for all four participants. Study 4 showed that a trial-based DRO increased all participants' latency to vocal stereotypy during DTT sessions relative to baseline. However, a clinically acceptable duration of 5 min or more was only obtained for the one participant who showed subsequent decreases in vocal stereotypy with the HP-LS stimulus in Study 3. Results of Study 5 showed that a DRA procedure in which participants earn access to the HP-LS item for correct responses on acquisition targets during DTT effectively decreased vocal stereotypy and increased correct responding for all participants. Taken together, the results of all five studies suggest that immediate and subsequent effects of stimuli on vocal stereotypy may be used to inform treatment decisions.

The findings of the current study are consistent with the findings of Frewing et al. (2015) in that we were able to identify a reliable HP-LS stimulus in as little as three 10-min sessions. However, the FOCSA was less accurate in predicting items that evoked stereotypy (i.e., HP-HS and LP-HS items) which is also consistent with the findings of Frewing et al. Further, the current data set showed consistency in item engagement at high levels for items predicted as high preference by the FOCSA and low levels for items predicted as low preference by the FOCSA. Data in the TCMS showed that three of our participants (Alex, Mark and Nick) engaged with the HP-LS at consistently higher percentages over the item that increased stereotypy. Sam, however, engaged with the HP-HS stimulus at consistently higher percentages than the HP-LS. These differences in item engagement across participiants are consistent with the results of Frewing et al. in that some participants exhibited higher preferrence for items that increased stereotypy and some exhibited higher preference for items that decreased stereotypy. Taken together, the results of the current study and Frewing et al. suggest that preference alone will not reliably identify a stimulus that decreases stereotypy. In other words, simply conducting a preference assessment will not consistently lead to the identification of stimuli that decrease engagement in stereotypy. Further, conducting a preference assessment alone may lead to the identification of a stimulus that is preferred but increases engagement in stereotypy. If a practitioner is attempting to identify a stimulus to use in the treatment of stereotypy, it is imperative that they take data in the way the FOCSA prescribes and calculate conditional probabilities in order to identify an item that decreases immediate levels of stereotypy and is high preference.

Results of Study 3 suggest that NCR is a viable treatment option when the FOCSA predicts a HP-LS stimulus. The TCMS validated the FOCSA predictions of the HP-LS in that all participants engaged in consistently lower percentages of vocal stereotypy during the HP-LS sequence relative to the control. Further, results suggest that not only was the HP-LS a structural match in that it produced the same sensory stimulation hypothesized to maintain vocal stereotypy (i.e., auditory stimulation), but it is also a functional match for all participants' vocal stereotypy. Rapp (2007) stated that a functionally matched stimulus is one that does not simply compete with stereotypy but rather one that replaces the reinforcement produced by stereotypy. Rapp (2007) suggested that a functionally matched stimulus is evidenced by lower levels of the target behavior both during NCR and lower or unchanged levels in the target behavior (compared to the NI sequence) after NCR is removed. Several published studies that employed NCR with functionally matched stimulation during a TCMS found that a decrease in subsequent stereotypy is rare (Rapp 2006, 2007; Rapp et al., 2013). The rarity of an immediate and subsequent decrease is consistent with the current study in that a subsequent decrease in targeted stereotypy was only seen in Sam. This data pattern seems to be rare in the literature but may be the best indicator of a functionally matched stimulus for stereotypy (Rapp et al., 2013).

Although the HP-LS stimulus in the current study was not matched to nontargeted stereotypy, all participants' nontargeted stereotypy decreased during the second component of the HP-LS sequence in the TCMS. Further, the HP-LS did not seem to exert subsequent effects on nontargeted stereotypy for any participant. These results are consistent with the results of Rapp et al. (2017) who found that nontargeted stereotypy decreased during noncontingent access to a stimulus that was functionally matched to targeted stereotypy. These desirable results on nontargeted stereotypy are noteworthy in that previous literature has reported increases in other topographies of stereotypy when targeted stereotypy is decreased (e.g., Rapp et al., 2004, 2013). Rapp et al. (2017) also reported that when a functionally matched stimulus was provided contingent upon the omission of targeted stereotypy during a DRO preparation, that latency to nontargeted stereotypy increased as well as latency to targeted stereotypy. Although we did not report nontargeted stereotypy during the DRO phase of the current experiment, we can hypothesize that similar effects would be achieved since similar data patterns in the TCMS were present. Addtionally, the decrease in nontargeted stereotypy for all participants in the DRA suggest that the HP-LS item may have decreased nontargeted stereotypy in the DRO as well.

Results of Study 4 suggest that a DRO procedure may only be a viable treatment option for individuals who show a subsequent decrease in stereotypy following NCR with the HP-LS stimulus. Although latency to vocal stereotypy increased relative to baseline for all participants, researchers were only able to reach a clinically acceptable duration of 5 min or more for the one participant (Sam) who showed a subsequent decrease in stereotypy as evidenced in component 3 of the TCMS. The other two participants (Mark and Alex) only reached durations of about 30 s and the clinical utility of these short durations is questionable (Rapp & Lanovaz, 2016). These results suggest that if a clinician is considering a DRO treatment to decrease vocal stereotypy, they should consider implementing a TCMS to evaluate what subsequent effects NCR with the HP-LS has on stereotypy. If a subsequent decrease is shown, DRO may be a viable treatment option. However, if no subsequent decrease is shown, DRO may not be the best treatment option and the clinician should consider either NCR with the HP-LS or DRA with the HP-LS as outlined in Study 5.

The results of Study 5 suggest that DRA with a HP-LS stimulus may be a useful treatment option that not only decreases stereotypy for inividuals regardless of subsequent effects on stereotypy but also targets alternative skills for increase. Additionally, this preparation may also decrease nontargeted forms of stereotypy for some participants. This procedure is easy to implement and not vastly different from typical DTT preparations. As such, it may be easy to train caregivers, teachers and direct line staff to implement these procedures. Alex exhibited similar decreases in stereotypy during the DRA with edibles condition. However, it is unlikely that edible items are matched to vocal stereotypy and the generality of a DRA with edibles to decrease vocal stereotypy is unknown given the current data set. Further, although vocal praise statements produce auditory stimulation, the DRA with praise procedure did not decrease Alex's vocal stereotypy. Similar to the DRA with edibles procedure, the generality of these effects are unknown.

The overall results of the current paper suggest that effective treatment options for vocal stereotypy may be derived from a FOCSA. The current data suggest that clinicians should conduct a FOCSA to identify a stimulus that can be used in treatment. Results from Study 2 suggest that high preference items may decrease or increase stereotypy so it is important to take data in the way that the FOCSA prescribes to ensure a HP-LS item is identified. Once a clinician implements a FOCSA and identifies a HP-LS stimulus, they have the treatment options of NCR, DRO and DRA with the HP-LS stimulus. The current data suggest that NCR and DRA may be effective across a variety of individuals regardless of subsequent effects. However, a clinician should consider implmenting a TCMS to identify subsequent effects when considering a DRO.

There are some limitations to the current studies. The first limitation is that the increases in DRO durations in Study 4 were somewhat arbitrary. We doubled initial durations but when durations were larger or when a participant failed to meet a duration, increases were more marginal. As such, it is unclear as to whether differences in duration increases may have led to increased success for some participants. The second limitation is that it is unclear if the results of Sam's DRO could have been reached with the use of an arbitrary reinforcer. The third limitation is that it is unclear if DRA with an arbitrary reinforcer could obtain the same results in Study 5. Researchers exposed Alex to a phase in which he earned edibles for correct responses in the DRA and we saw comparable decreases in vocal stereotypy to a DRA phase in which he earned the HP-LS for correct responses. As such, it is unclear as to whether or not we could have obtained similar results for the other participants. However, their baseline IRT in the DRO phase suggest that providing edibles for correct responses during DTT would not have effectively decreased their vocal stereotypy to the same levels as the HP-LS. During non-experimental sessions, Mark rarely accepted edibles and Nick only accepted edibles about half the time.

Future research should evaluate the most effective method by which to thin DRO durations. Researchers may consider employing a 10% increase of the previous duration or a percentile schedule to see if these methods lead to more successful omission of the target behavior. Future research should also evaluate whether an arbitrary reinforcer such as an edible may be used in the DRO or DRA procedures as effectively as the HP-LS stimulus. As previously noted, DRA with edibles successfully decreased targeted stereotypy for Alex however subsequent effects of the procedure are unknown. As such, the efficacy of arbitrary reinforcers in a DRO or DRA preparation should be evaluated with a TCMS. Lastly, the current paper suggests that there is still more research to be done in relation to what levels of increased conditional percentages of stereotypy should be labeled as a “high stereotypy” item. The current data suggest that no high stereotypy item reliably increased stereotypy above baseline levels. As such, future research should be directed at what exactly constitutes a high stereotypy item and whether or not this item can be used in the treatment of stereotypy.

References

Cook, J. L., Rapp, J. T., Gomes, L. A., Frazer, T. J., & Lindblad, T. L. (2014). Effects of verbal reprimands on targeted and untargeted stereotypy. *Behavioral Interventions*, *29*(2), 106-124. doi: 10.1002/bin.1378

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis.* Upper Saddle River, NJ: Pearson Education, Inc.

Cunningham, A. B., & Schreibman, L. (2008). Stereotypy in autism: The importance of function. *Research in Autism Spectrum Disorders*, *2*, 469-479. doi: 0.1016/j.rasd.2007.09.006

Devine, S. L., Rapp, J. T., Testa, J. R., Henrickson, M. L., & Schnerch, G. (2011). Detecting changes in simulated events using partial-interval recording and momentary time sampling III: Evaluating sensitivity as a function of session length. *Behavioral Interventions, 24*, 103-124. doi: 10.1002/bin.328

DiGennaro Reed, F. D. D., Hirst, J. M., & Hyman, S. R. (2012). Assessment and treatment of stereotypic behavior in children with autism and other developmental disabilities: A thirty year review. *Research in Autism Spectrum Disorders*, *6*, 422-430. doi: 10.1016/j.rasd.2011.07.003

Enloe, K. A., & Rapp, J. T. (2014). Effects of noncontingent social interaction on immediate and subsequent engagement in vocal and motor stereotypy in children with autism. *Behavior Modification*, *38*, 374-91. doi: 10.1177/0145445513514081

Frewing, T. M., Rapp, J. T., & Pastrana, S. J. (2015). Using conditional probabilities during free-operant stimulus preference assessments to predict the effects of preferred items on stereotypy: Preliminary findings. *Behavior Modification, 39*, 740–765. doi: 10.1177/0145445515593511

Frewing, T. M., Tanner, A., Bonner, A. C., Baxter, S., & Pastrana, S. J. (2015). The immediate and subsequent effects of response interruption and redirection on vocal stereotypy, motor stereotypy and heart rate. *European Journal of Behavior Analysis*, *16*(2), 351-361. doi: 10.1080/15021149.2015.1110407

Gardenier, N. C., MacDonald, R., & Green, G. (2004). Comparison of direct observational methods for measuring stereotypic behavior in children with autism spectrum disorders. *Research in Developmental Disabilities*, *25*(2), 99-118. doi:10.1016/j.ridd.2003.05.004

Gehrman C., Wilder D. A., Forton A. .P, & Albert, K. (2017). Comparing resetting to non-resetting DRO procedures to reduce stereotypy in a child with autism. *Behavioral Interventions*. Advance online publication. doi: 10.1002/bin.1486

Harrop, C., McBee, M., & Boyd, B. A. (2016). How are child restricted and repetitive behaviors associated with caregiver stress over time? A parallel process multilevel growth model. *Journal of Autism and Developmental Disorders*. Advance online publication. doi: 10.1007/s10803-016-2707-7

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*, 197–209. doi: 10.1901/jaba.1994.27-197 (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2,* 3–20, 1982)

Kazdin, A. E. (2011). Single-case research designs: *Methods for clinical and applied settings*. Oxford University Press.

Koegel, R. L., Firestone, P. B., Kramme, K. W., & Dunlap, G. (1974). Increasing spontaneous play by suppressing self-stimulation in autistic children. *Journal of Applied Behavior Analysis*, *7*(4), 521-528. doi: 10.1901/jaba.1974.7-521

Lanovaz, M. J., & Argumedes, M. (2010). Immediate and subsequent effects of differential reinforcement of other behavior and noncontingent matched stimulation on stereotypy. *Behavioral Interventions*, *25*(3), 229-238. doi: 10.1002/bin.308

Lanovaz, M. J., Fletcher, S. M., & Rapp, J. T. (2009). Identifying stimuli that alter immediate and subsequent levels of vocal stereotypy: A further analysis of functionally matched stimulation. *Behavior Modification, 33*, 682-704. doi: 10.1177/0145445509344972

Lanovaz, M. J., Rapp, J. T., & Fletcher, S. E. (2010). Expanding functional analysis of automatically reinforced behavior using a three-component multiple-schedule. *European Journal of Behavior Analysis*, *11*, 17-27. doi: 10.1080/15021149.2010.11434331

Lanovaz, M. J., Rapp, J. T., Long, E. S., Richling, S. M., & Carroll, R. A. (2014). Preliminary effects of conditioned motivating operations on stereotypy. *The Psychological Record*, *64*, 209–216. doi: 10.1007/s40732-014-0027-x

Lanovaz, M. J., Rapp, J. T., Maciw, I., Prégent-Pelletier, É., Dorion, C., Ferguson, S., & Saade, S. (2014). Effects of multiple interventions for reducing vocal stereotypy: Developing a sequential intervention model. *Research in Autism Spectrum Disorders, 8*, 529-545. doi: 10.1016/j.rasd.2014.01.009

Lanovaz, M. J., Robertson, K. M., Soerono, K., & Watkins, N. (2013). Effects of reducing stereotypy on other behaviors: A systematic review. *Research in Autism Spectrum Disorders*, *7*(10), 1234-1243. doi:10.1016/j.rasd.2013.07.009

Lanovaz, M. J., & Sladeczek, I. E. (2012). Vocal stereotypy in individuals with autism spectrum disorders a review of behavioral interventions. *Behavior Modification*, *36*, 146-164. doi: 10.1177/0145445511427192

Lanovaz, M. J., Sladeczek, I. E., & Rapp, J. T. (2012). Effects of noncontingent music on vocal stereotypy and toy manipulation in children with autism spectrum disorders. *Behavioral Interventions*, *27*, 207-223. doi: 10.1002/bin.1345

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis*, *36*, 407-414. doi 10.1901/jaba.2003.36-407

Meany-Daboul, M. G., Roscoe, E. M., Bourret, J. C., & Ahearn, W. H. (2007). A comparison of momentary time sampling and partial-interval recording for evaluating functional relations. *Journal of Applied Behavior Analysis*, *40*(3), 501-514. doi: 10.1901/jaba.2007.40-501

Morrison, H., Roscoe, E. M., & Atwell, A. (2011). An evaluation of antecedent exercise on behavior maintained by automatic reinforcement using a three-component multiple schedule. *Journal of applied behavior analysis*, *44*(3), 523-541. doi: 10.1901/jaba.2011.44-523

Mudford, O. C., Taylor, S. A., & Martin, N. T. (2009). Continuous recording and interobserver agreement algorithms reported in the Journal of Applied Behavior Analysis (1995–2005). *Journal of Applied Behavior Analysis*, *42*(1), 165-169. doi: 10.1901/jaba.2009.42-165

Pastrana, S. J., Rapp, J. T., & Frewing, T. M. (2013). Immediate and subsequent effects of response interruption and redirection on targeted and untargeted forms of stereotypy. *Behavior Modification*, *37*, 591-610. doi: 10.1177/0145445513485751

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*(1), 13-27. doi: 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*(1), 47-60. doi: 10.1002/jaba.26

Rapp, J. T., (2004). Effects of prior access and environmental enrichment on stereotypy. *Behavioral Interventions*,*19*, 287-295. doi: 10.1002/bin.166

Rapp, J. T. (2005). Some effects of audio and visual stimulation on multiple forms of stereotypy. *Behavioral Interventions*, *20*, 255-272. doi: 10.1002/bin.196

Rapp, J. T. (2006). Toward an empirical method for identifying matched stimulation for automatically reinforced behavior: A preliminary investigation. *Journal of Applied Behavior Analysis*, *39*(1), 137-140. doi: 10.1901/jaba.2006.37-05

Rapp, J. T. (2007). Further evaluation of methods for identifying matched stimulation. *Journal of Applied Behavior Analysis*, *40*, 73-88. doi: 10.1901/jaba.2007.142-05

Rapp, J. T. (2008). Conjugate reinforcement: A brief review and suggestions for applications to the assessment of automatically reinforced behavior. *Behavioral Interventions*, *23*, 113–136. doi: 10.1002/bin.259

Rapp, J. T., Colby, A. M., Vollmer, T. R., Roane, H. S., Lomaas, J., & Britton, L. M. (2007). Interval recording for duration events: An extensive re-evaluation. *Behavioral Interventions*, *22*, 319-345. doi: 10.1002/bin.239

Rapp, J. T., Colby-Dirksen, A. M., Michalski, D. N., Carroll, R. A., & Lindenberg, A. M. (2008). Detecting changes in simulated events using partial-interval recording and momentary time sampling. *Behavioral Interventions*, *23*, 237-269. doi: 10.1002/bin.269

Rapp, J. T., Cook, J. L., McHugh, M., & Mann, K. (2017). Decreasing stereotypy using NCR and DRO with functionally matched stimulation: Effects on targeted and nontargeted stereotypy. *Behavior Modification*, *41*, 45–83. doi: 10.1177/0145445516652370

Rapp, J. T., & Lanovaz, M. J. (2016). *Evidence-Based Practices in Behavioral Health: Stereotypy*. In N. Singh (Ed.), Clinical Handbook of Evidence-Based Practices for Individuals with Intellectual Disabilities. Springer.

Rapp, J. T., Swanson, G. S., Sheridan, S., Enloe, K., Maltese, D., Sennott, L., Lanovaz, M. J. (2013). Immediate and subsequent effects of matched and unmatched stimuli on targeted vocal stereotypy and untargeted motor stereotypy. *Behavior Modification*, *37*, 543-567. doi: 10.1177/0145445512461650.

Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities*, *26*(6), 527-547. doi:10.1016/j.ridd.2004.11.005

Rapp, J. T., Vollmer, T. R., Peter, C., Dozier, C. L., & Cotnoir, N. M. (2004). Analysis of response allocation in individuals with multiple forms of stereotyped behavior. *Journal of Applied Behavior Analysis*, *37*(4), 481-501. doi: 10.1901/jaba.2004.37-481

Richman, D. M., Barnard-Brak, L., Grubb, L., Bosch, A., & Abby, L. (2015). Meta-analysis of noncontingent reinforcement effects on problem behavior. *Journal of Applied Behavior Analysis, 48(1),* 131-152. doi: 10.1002/jaba.189

Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, *31*(4), 605-620. doi: 10.1901/jaba.1998.31-605

Van Camp, C. M., Lerman, D. C., Kelley, M. E., Roane, H. S., Contrucci, S. A., & Vorndran, C. M. (2000). Further analysis of idiosyncratic antecedent influences during the assessment and treatment of problem behavior. *Journal of Applied Behavior Analysis*, *33*(2), 207-221. doi: 10.1901/jaba.2000.33-207

Vollmer, T. R., & Iwata, B. A. (1992). Differential reinforcement as treatment for behavior disorders: Procedural and functional variations. *Research in Developmental Disabilities*, *13*(4), 393-417.

Vollmer, T. R., Marcus, B. A., Ringdahl, J. E., & Roane, H. S. (1995). Progressing from brief assessments to extended experimental analyses in the evaluation of aberrant behavior. *Journal of Applied Behavior Analysis*, *28*, 561-576.

Table 1

Response Definitions for Targeted and Nontargeted Behavior

Participant	Targeted Response	Nontargeted Response
Mark	***Vocal stereotypy:*** Any repetitive or invariant noncontextual or babbling vocalizations, at least 2, within 3 seconds or lasting longer than 3 seconds.	***Motor stereotypy:*** Any repetitive or invariant gross or fine motor movements, 2 or more repetitive movements within 3 seconds or lasting longer than 3 seconds. Includes finger movements.
Nick	***Vocal stereotypy*:** Any repetitive, noncontextual or babbling vocalizations, at least 2, within 2 s or lasting longer than 2 s.	***Mirror viewing*:** Any instance of eye gaze towards a mirror for 2 s or more or twice within 2 s. Mirror viewing is mutually exclusive from any stereotypy that is reflected in the mirror (e.g., grimacing, arm posturing) but not from stereotypy that is not reflected in the mirror (e.g., vocal stereotypy or stomping) ***Stomping:*** Any instance of forceful foot to surface contact. Two or more instances within 2 s. ***Pacing:*** Continuous forward or backward walking movement for 2 s or longer. ***Posturing*:** Static leg or feet posturing (standing on tip toes) for 2 s or longer. ***Facial expressions:*** Squinting eyes, smiling, and puckering lips for 2 s or more when these expressions are not directed toward another person
Sam	***Vocal stereotypy:*** Any instance in which SPa emits repetitive (at least 2 instances within 2-s) or invariant (lasting at least 2-s) vocalizations. 2-s onset, 2-s offset.	***Repetitive Facial:*** Any repetitive (minimum of 2 instances within 2 s) lip movements such as repetitive puckering (excluding chewing), blowing/sucking air with rigid or stationary lips, repetitive blinking

		(minimum of two times within 2 s) or closing/squinting eyes for more than 2 s (excluding sleeping). ***Bruxing*:** clenching and grinding teeth togeether, resulting in an audible noise. Score with immediate onset and 3-s offset. ***Pacing***: Continuous forward or backward walking movement for 2-s or longer. ***Hand***: Repetitive (minimum of 2 times within 2-s) or invariant (lasting at least 2-s) movements with hands or fingers (without objects) ***Other Motor***: Any repetitive (minimum of 2 times within 2-s) or invariant (lasting at least 2-s) motor movements, including (but not limited to) rapid head shaking (moving head laterally from side to side within 1-s), spinning, waving of arms or hands
Alex	***Vocal stereotypy:*** Any repetitive, noncontextual or babbling vocalizations, at least 2, within 2 s or lasting longer than 2 s.	***Pacing:*** Continuous forward or backward walking movement for 2 s or longer. ***Hand**:* Repetitive (minimum of 2 times within 2-s) or invariant (lasting at least 2-s) movements with hands or fingers (without objects). Includes hand flapping, posturing the pointer finger and repetitive movements with fingers. ***Surface rubbing:*** Any forward, backward and/or sideways movement of the hand while making hand to surface (i.e., table, wall, floor) contact for 2-s or more. Also included is any repeated (2 times within 2-s) brief contact of surfaces to hand.

Table 2

Stimulus Classifications Resulting from the FOCSA

Participant	Stimulus	Classification
Mark	iPad™	HP-LS
	Ship	HP-HS
Nick	iPad™	HP-LS
	Trains	LP-HS
Sam	iPad™	HP-LS
	Trains	HP-HS
Alex	iPad™	HP-LS
	Octonauts™	HP-HS

Note. HP-LS = High preference low stereotypy; HP-HS = High preference high stereotypy; LP-HS = Low preference high stereotypy

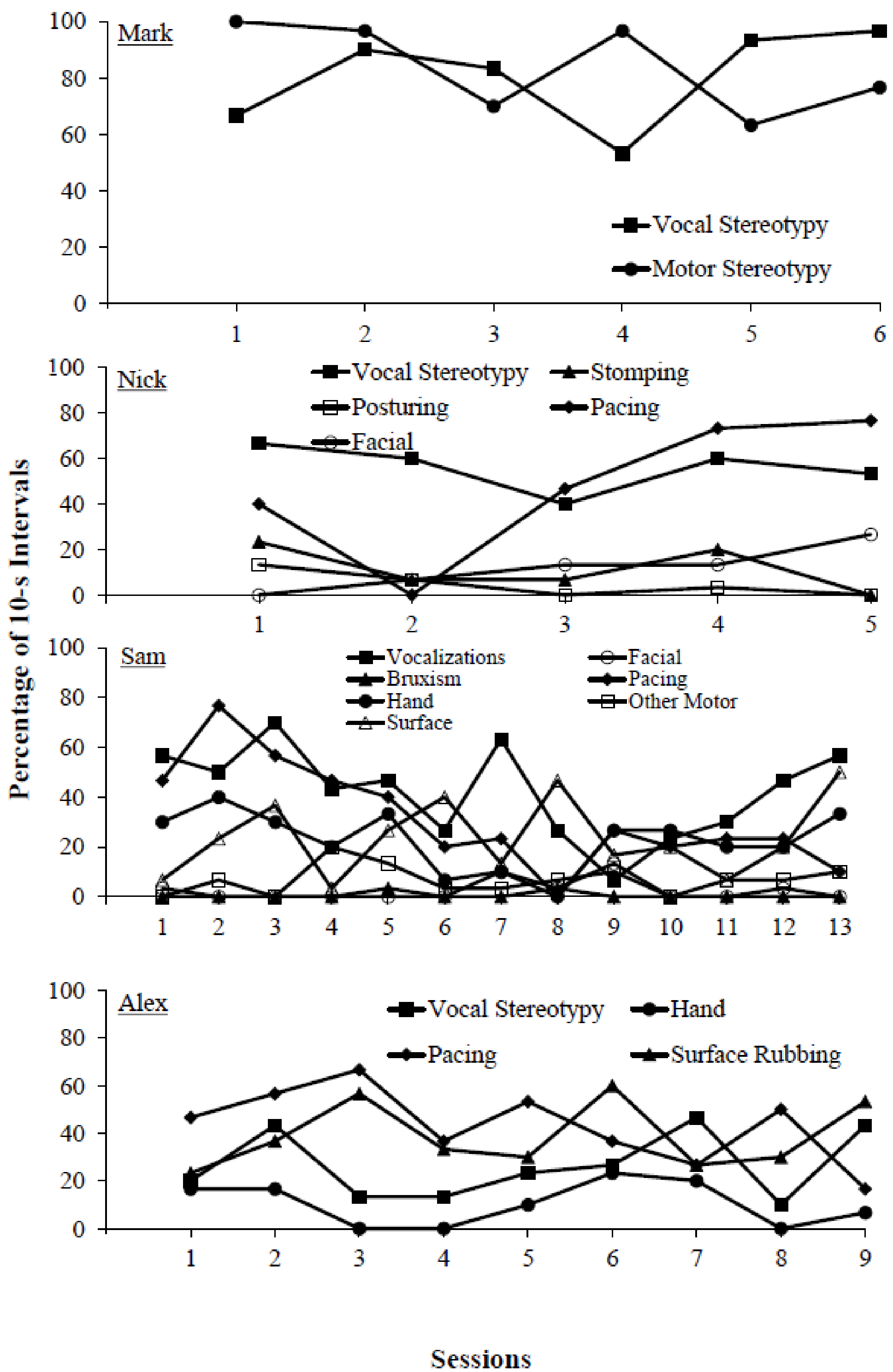

Figure 1. Percentage of 10-s intervals Mark (first panel), Nick (second panel), Sam (third panel), and Alex (fourth panel) engaged in repetitive behavior across consecutive no-interaction sessions.

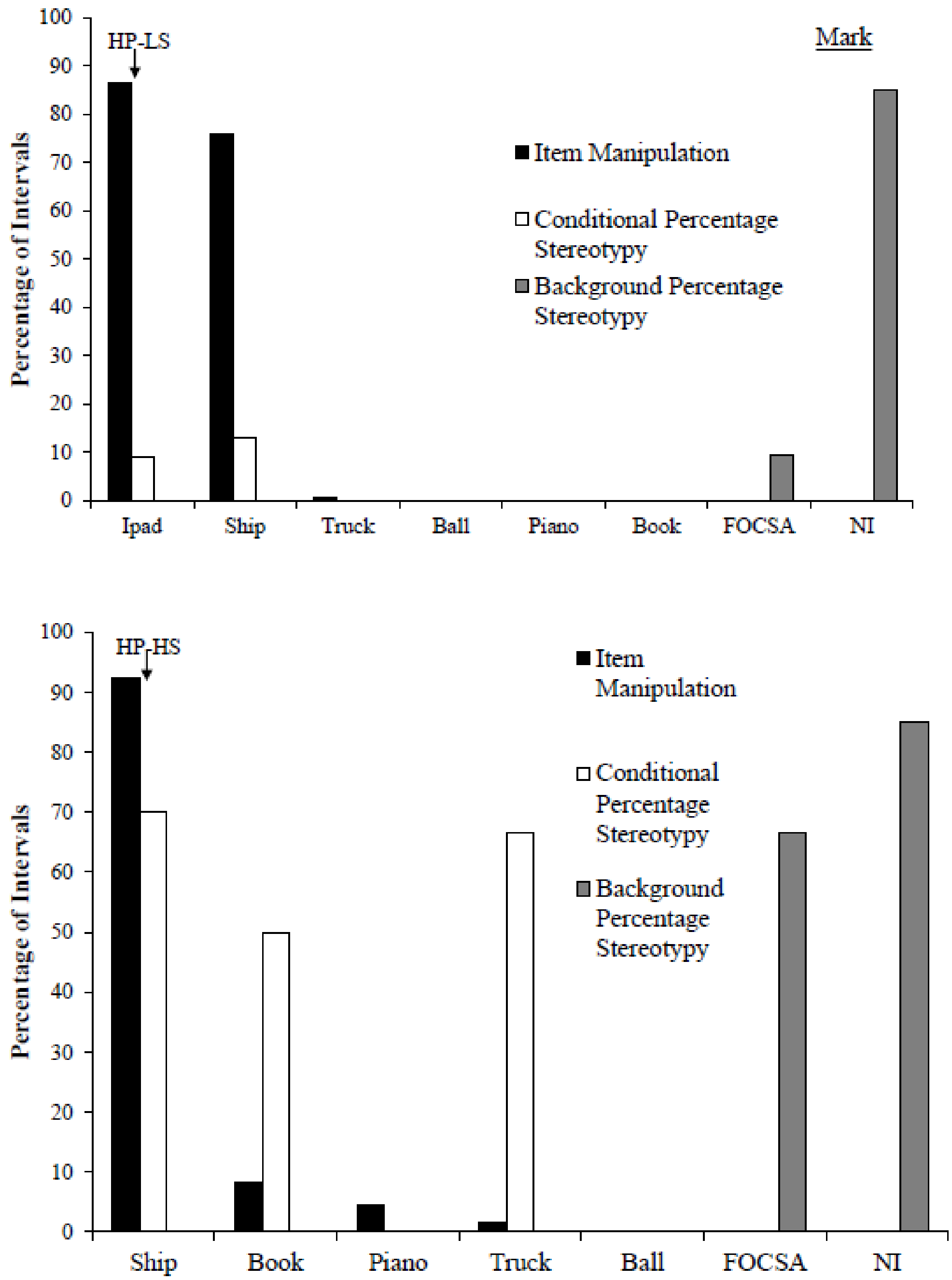

Figure 2. Percentage of 10-s intervals with item engagement, conditional percentage of stereotypy given item engagement and background percentage of stereotypy with iPad™ available (upper panel) and iPad™ restricted (lower panel) for Mark.

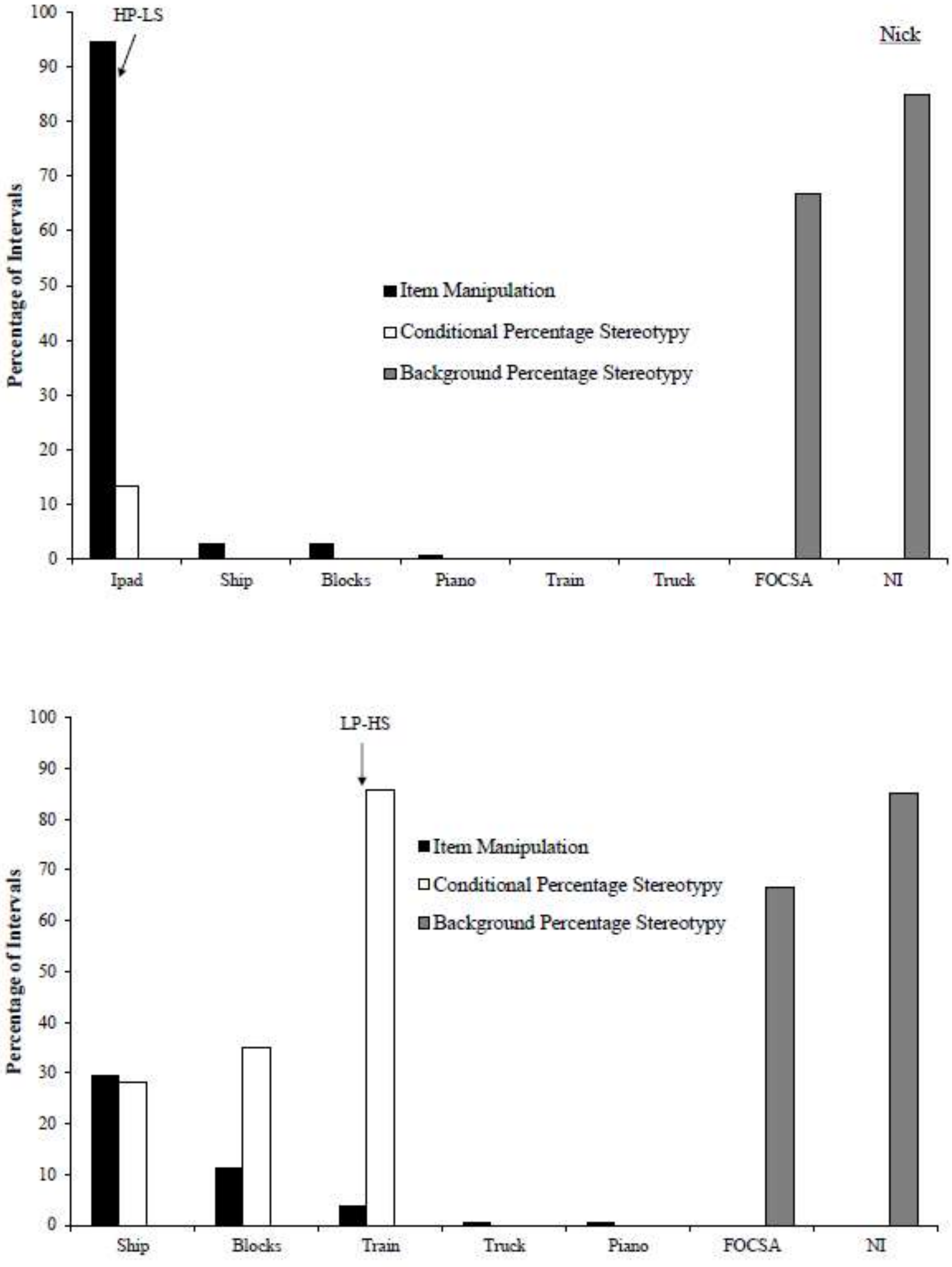

Figure 3. Percentage of 10-s intervals with item engagement, conditional percentage of stereotypy given item engagement and background percentage of stereotypy with iPad™ available (upper panel) and iPad™ restricted (lower panel) for Nick.

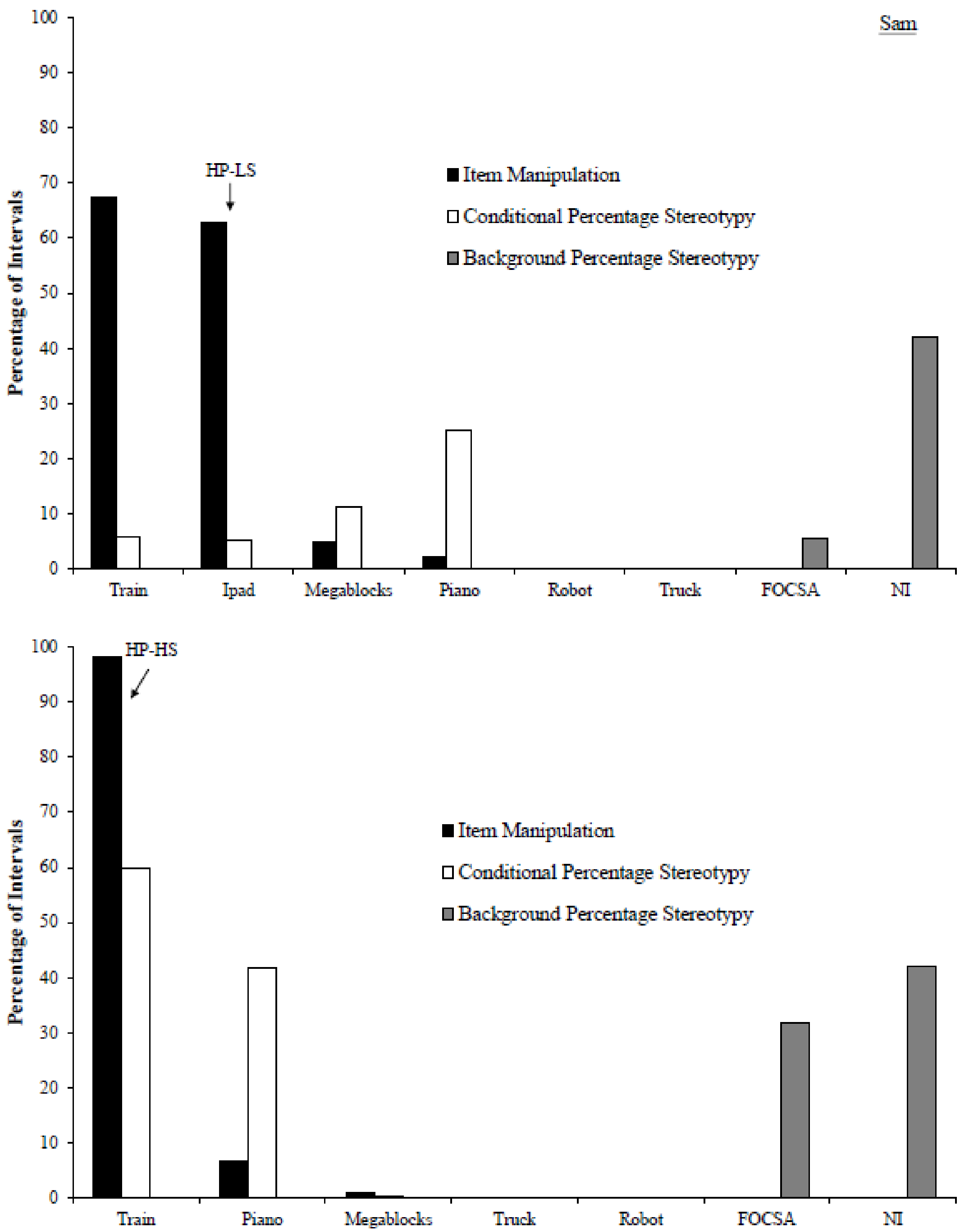

Figure 4. Percentage of 10-s intervals with item engagement, conditional percentage of stereotypy given item engagement and background percentage of stereotypy with iPad™ available (upper panel) and iPad™ restricted (lower panel) for Sam.

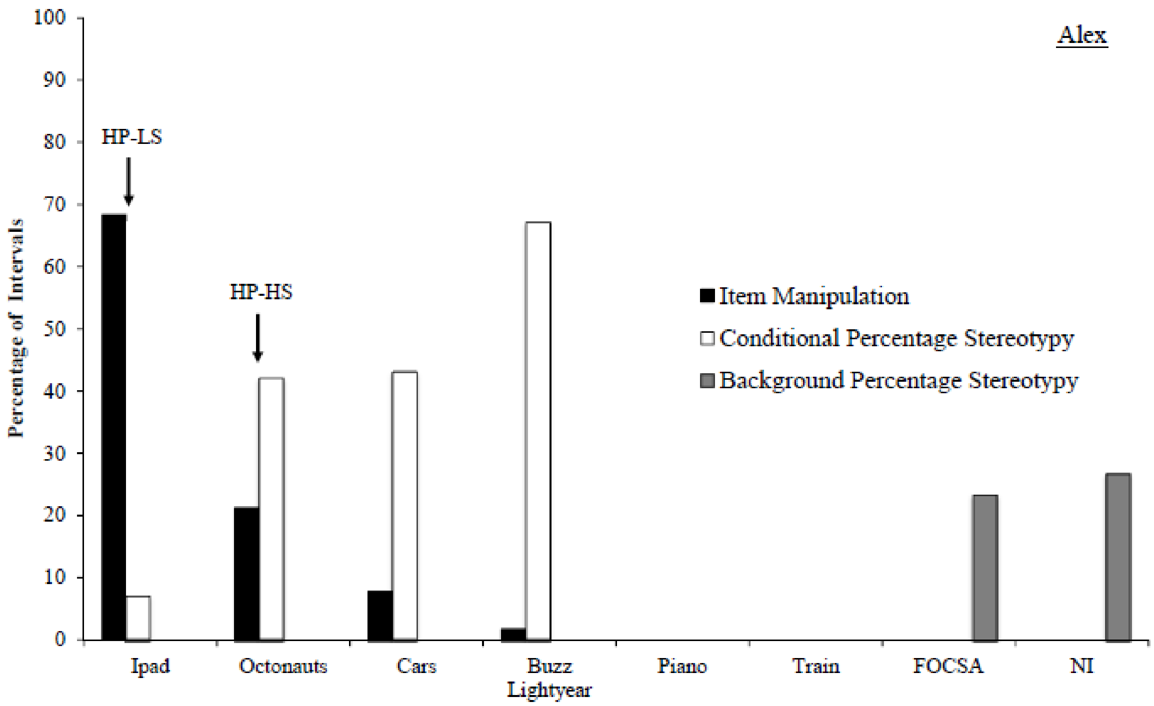

Figure 5. Percentage of 10-s intervals with item engagement, conditional percentage of stereotypy given item engagement and background percentage of stereotypy for Alex.

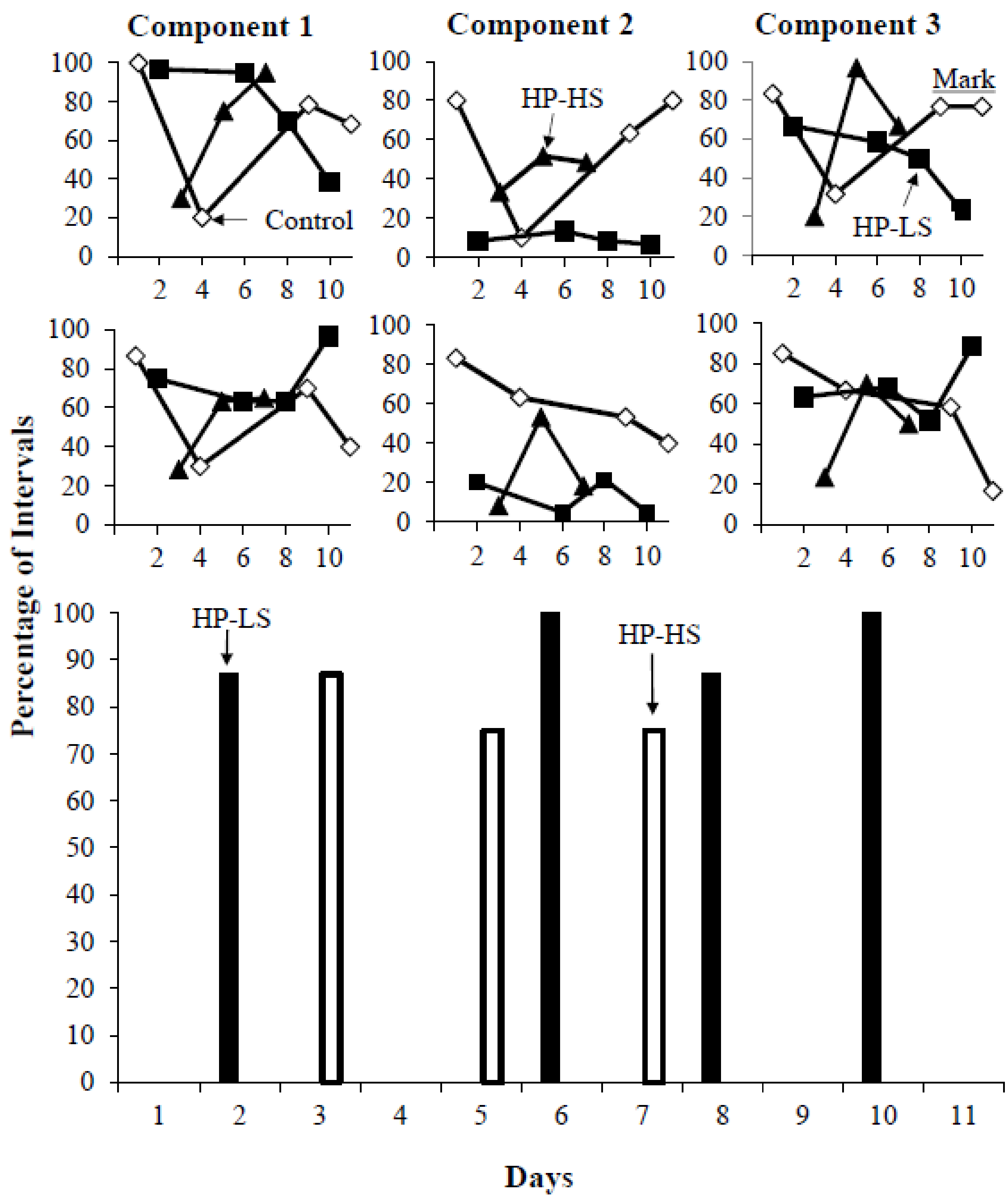

Figure 6. Percentage of 10-s intervals with vocal stereotypy (top panel) and motor stereotypy (middle panel) and item engagement during the second component of the HP-LS and HP-HS sequences (bottom panel) during the TCMS for Mark.

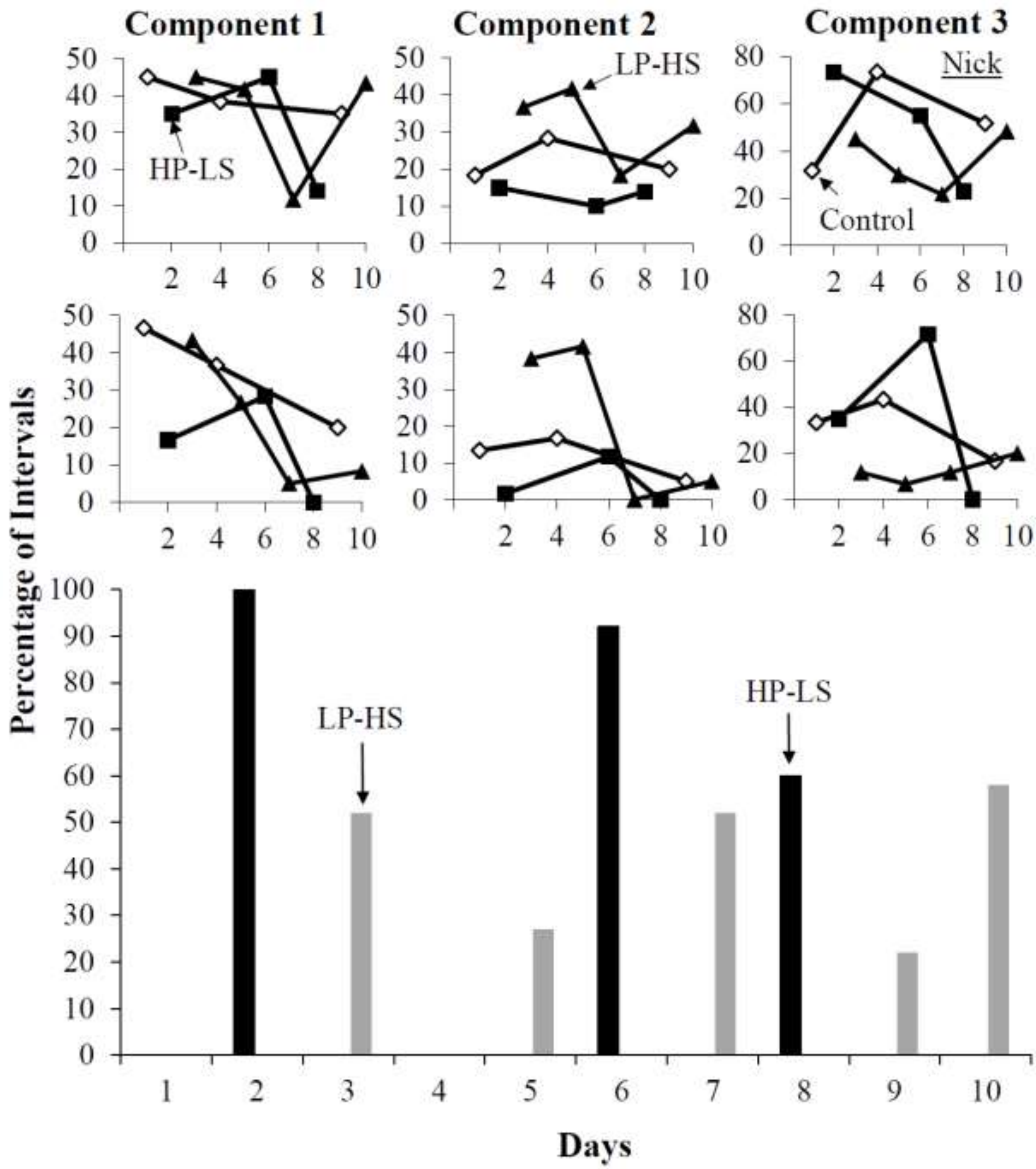

Figure 7. Percentage of 10-s intervals with vocal stereotypy (top panel) and motor stereotypy (middle panel) and item engagement during the second component of the HP-LS and HP-HS sequences (bottom panel) during the TCMS for Nick.

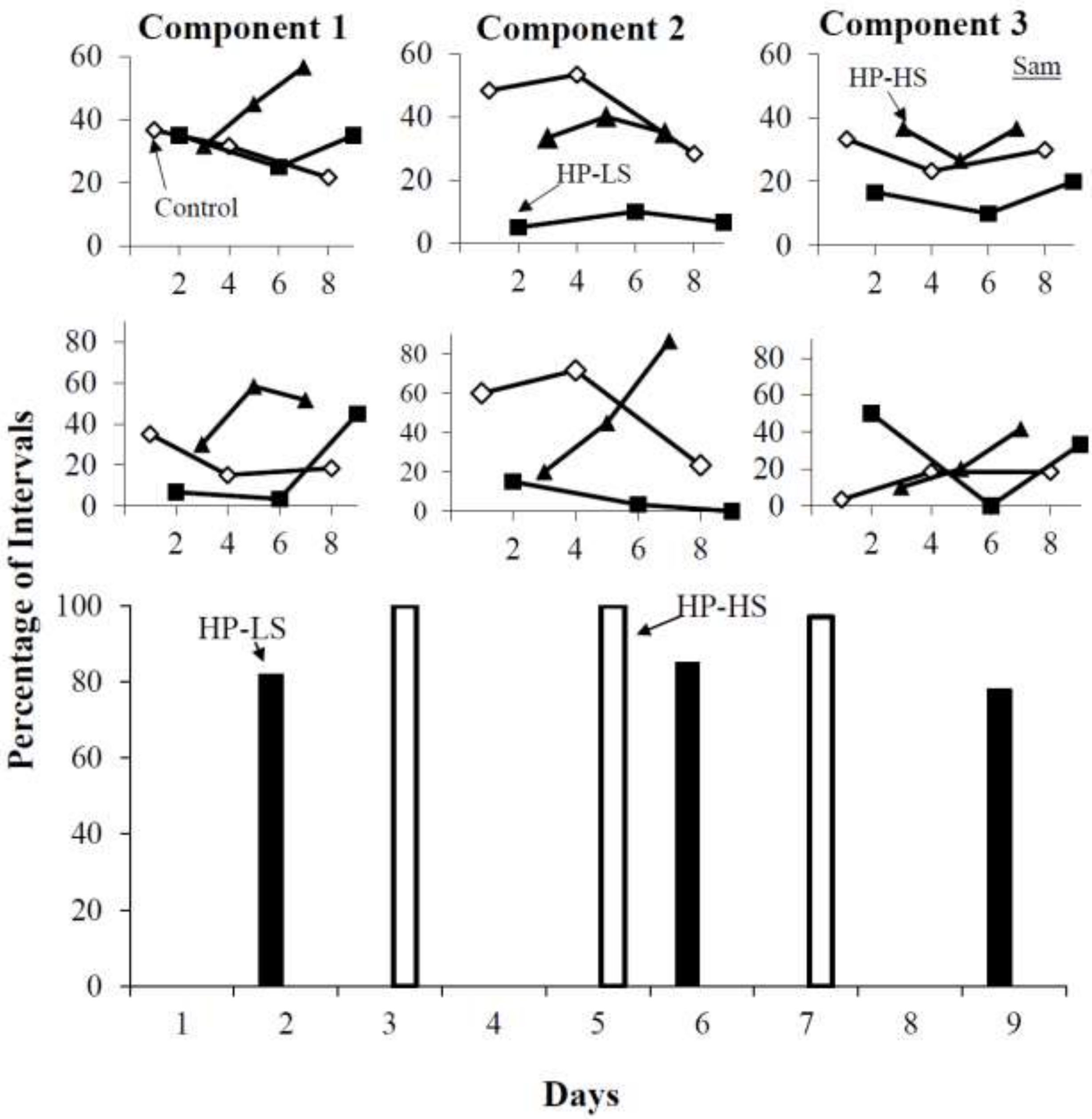

Figure 8. Percentage of 10-s intervals with vocal stereotypy (top panel) and motor stereotypy (middle panel) and item engagement during the second component of the HP-LS and HP-HS sequences (bottom panel) during the TCMS for Sam.

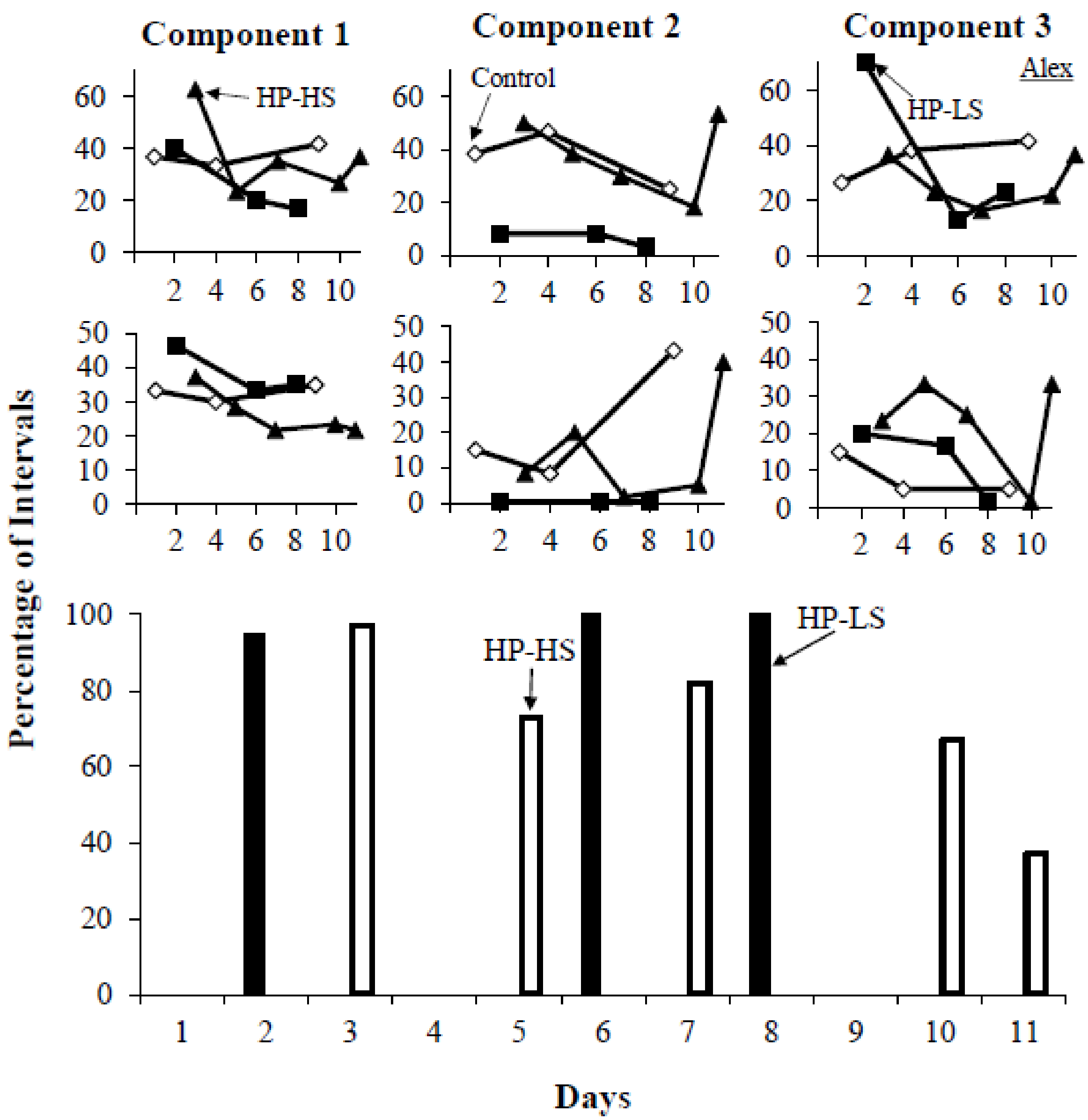

Figure 9. Percentage of 10-s intervals with vocal stereotypy (top panel) and motor stereotypy (middle panel) and item engagement during the second component of the HP-LS and HP-HS sequences (bottom panel) during the TCMS for Alex.

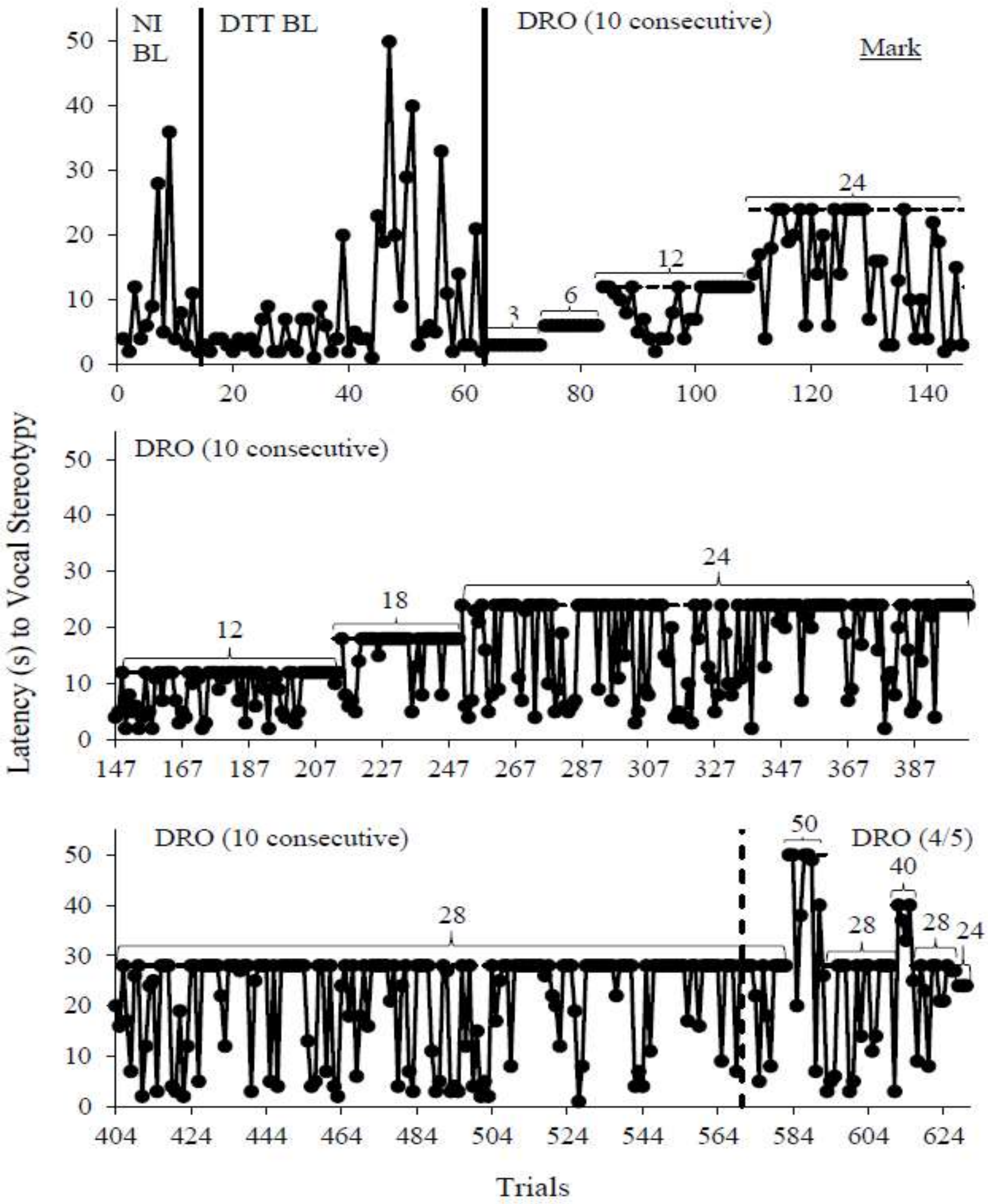

Figure 10. Latency in seconds to vocal stereotypy across trials 1 through 146 (upper panel), trials 147 through 403 (middle panel), and trials 404 through 625 (lower panle) during baseline and DRO phases for Mark. Numbers above the brackets denote the criterion in seconds.

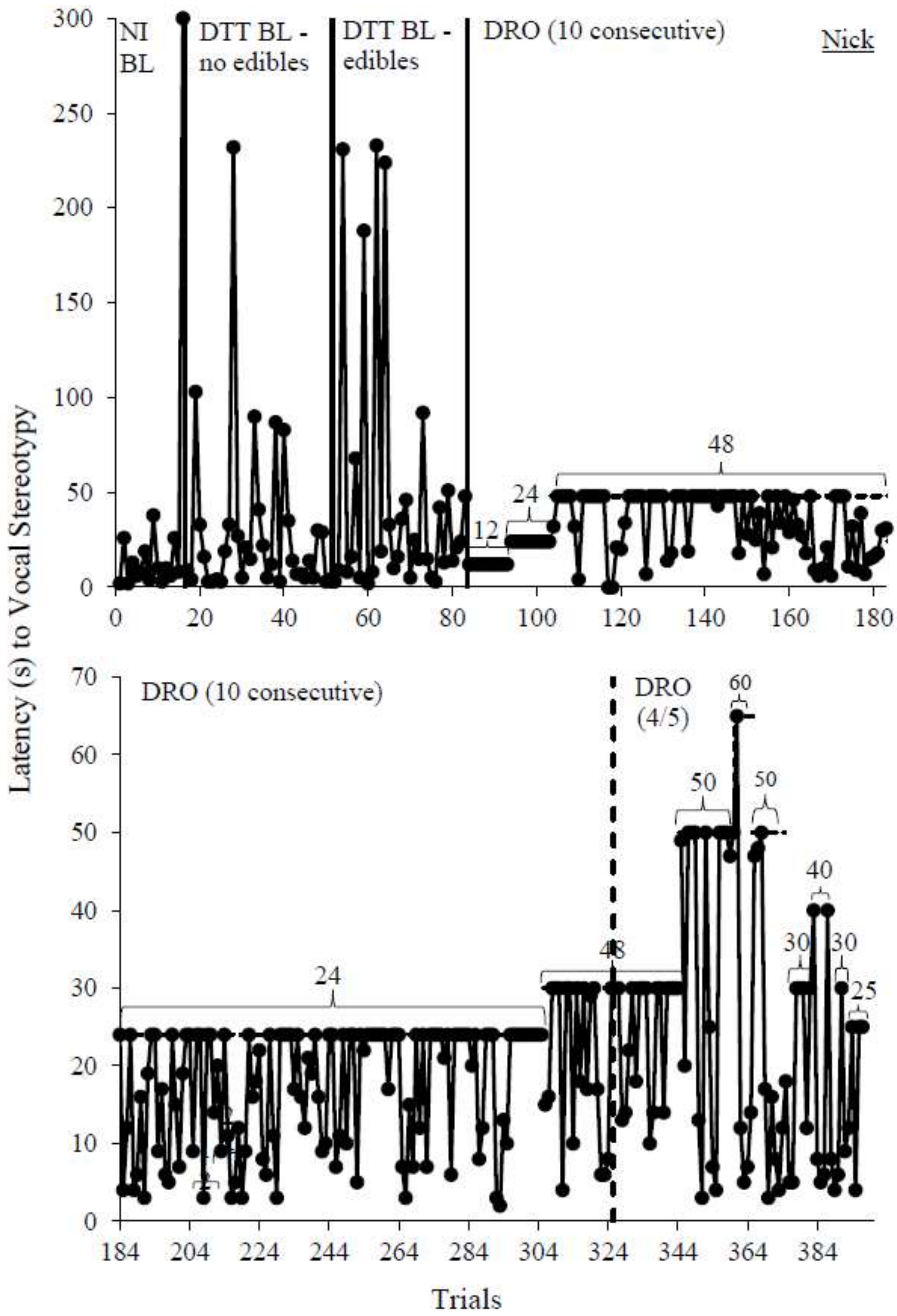

Figure 11. Latency in seconds to vocal stereotypy across trials 1 through 183 (upper panel), and trials 184 through 397 (lower panel) during baseline and DRO phases for Nick. Numbers above the brackets denote the criterion in seconds.

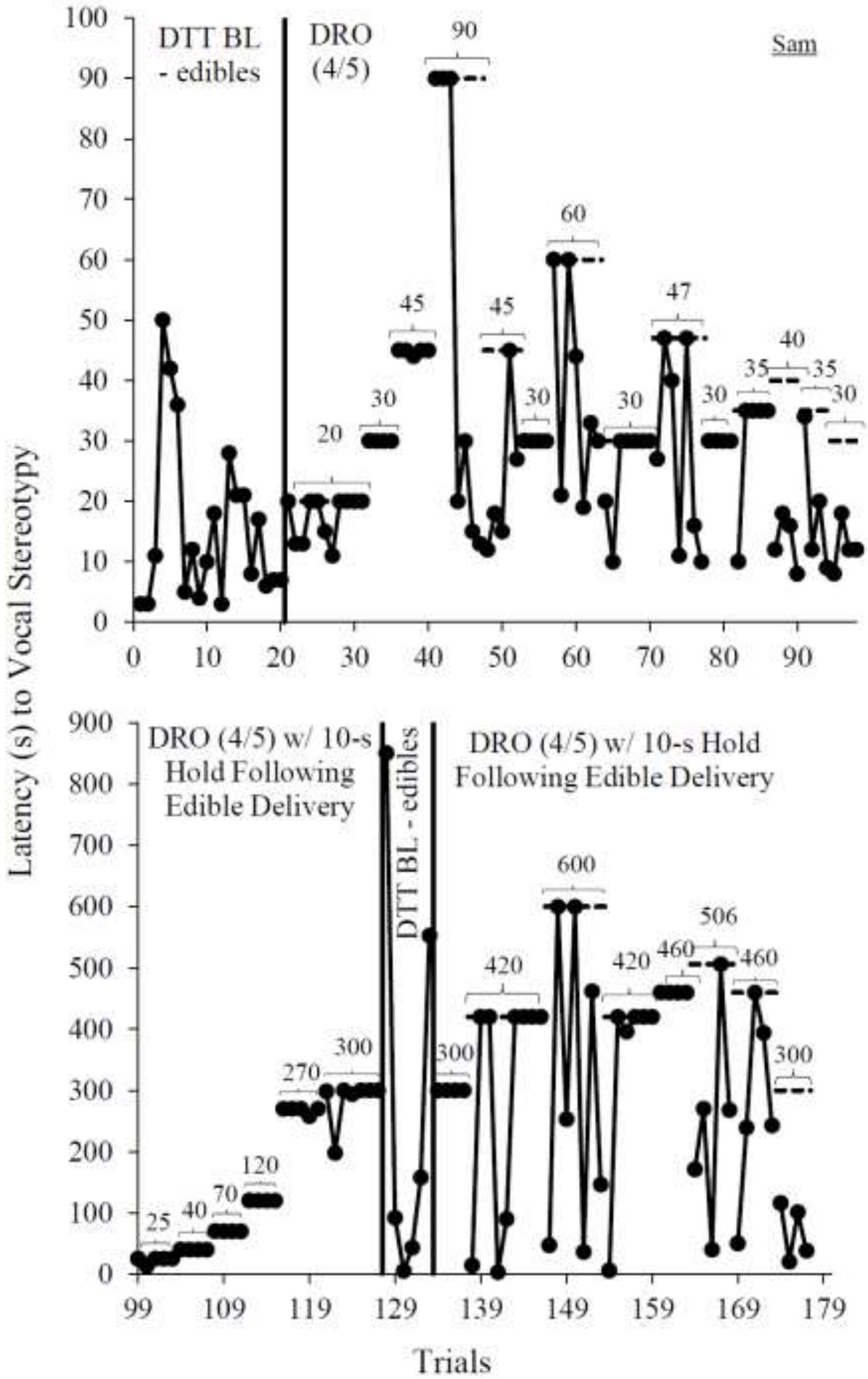

Figure 12. Latency in seconds to vocal stereotypy across trials 1 through 98 (upper panel), and trials 99 through 177 (lower panel) during baseline and DRO phases for Sam. Numbers above the brackets denote the criterion in seconds.

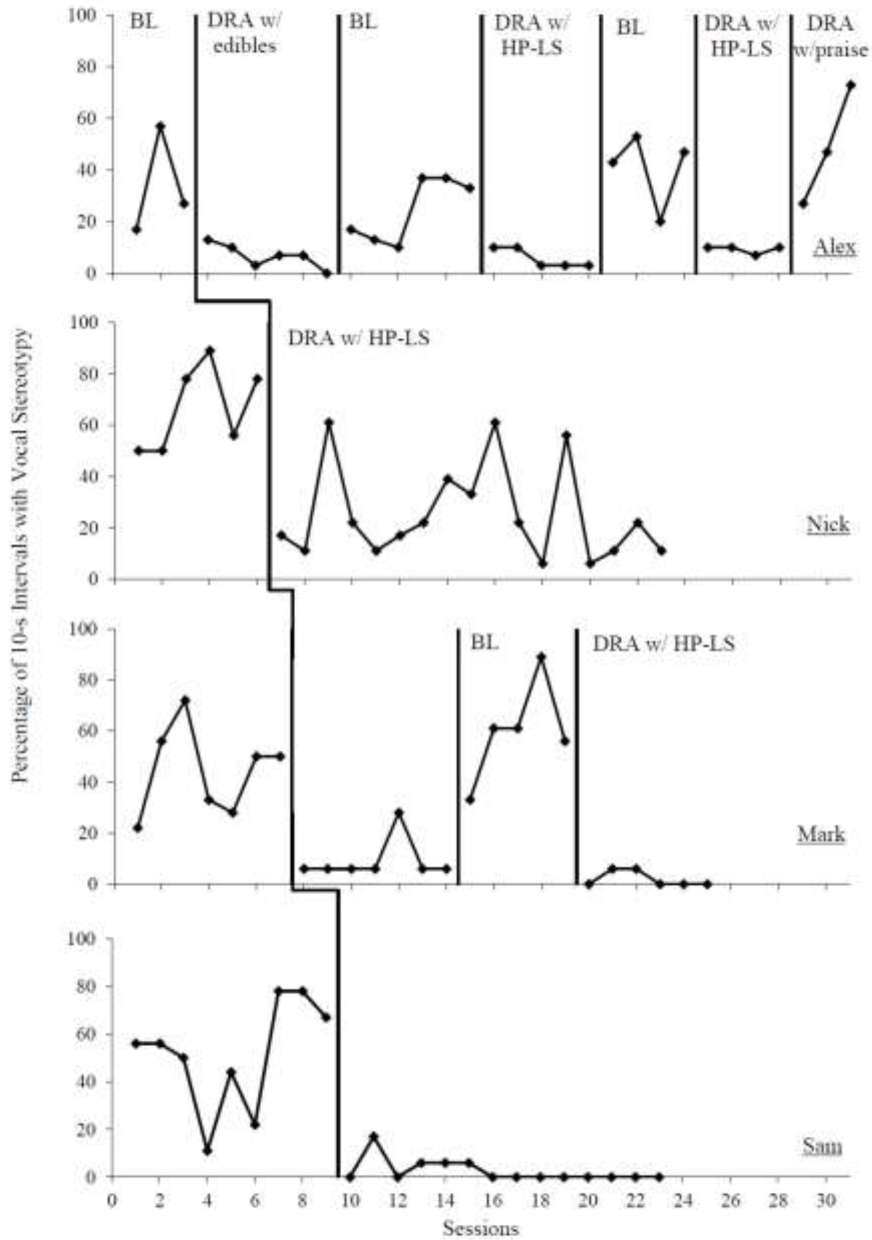

Figure 13. Percentage of 10-s intervals with vocal stereotypy across sessions throughout baseline (BL), differential reinforcement (DRA) with edibles, and DRA with high preference low stereotypy (HP-LS) phases for Alex (top panel), Nick (second panel), Mark (third panel) and Sam (bottom panel).

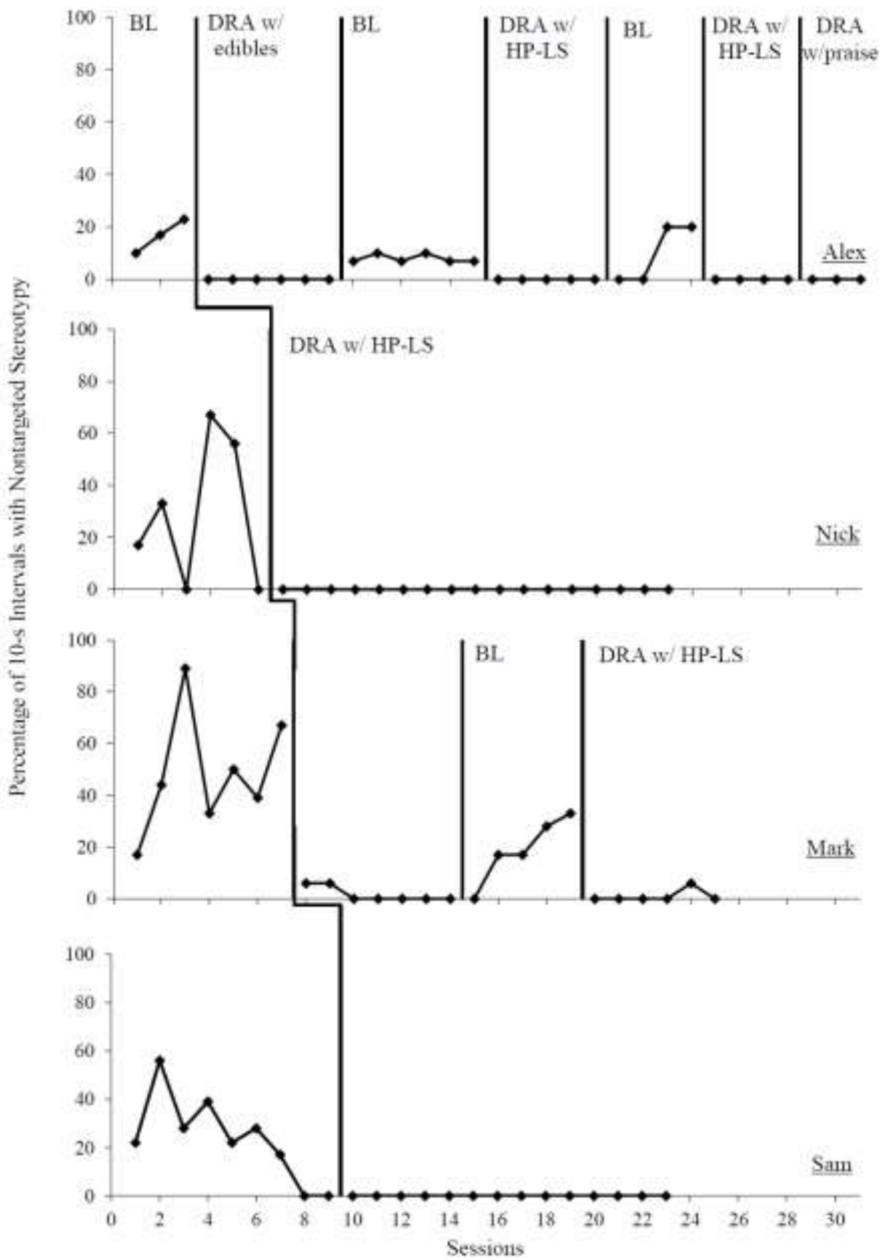

Figure 14. Percentage of 10-s intervals with nontargeted stereotypy across sessions throughout baseline (BL), differential reinforcement (DRA) with edibles, and DRA with high preference low stereotypy (HP-LS) phases for Alex (top panel), Nick (second panel), Mark (third panel) and Sam (bottom panel).