Behavior Modification
http://bmo.sagepub.com/

Identifying Stimuli that Alter Immediate and Subsequent Levels of Vocal Stereotypy: A Further Analysis of Functionally Matched Stimulation

Marc J. Lanovaz, Sarah E. Fletcher and John T. Rapp

Behav Modif 2009 33: 682
DOI: 10.1177/0145445509344972

Behavior Modification
Volume 33 Number 5
September 2009 682-704

10.1177/0145445509344972
http://bmo.sagepub.com
hosted at
http://online.sagepub.com

Identifying Stimuli that Alter Immediate and Subsequent Levels of Vocal Stereotypy

A Further Analysis of Functionally Matched Stimulation

Marc J. Lanovaz
Centre de Réadaptation Lisette-Dupras
Sarah E. Fletcher
St. Cloud State University
John T. Rapp
UK Young Autism Project

We used a three-component multiple-schedule with a brief reversal design to evaluate the effects of structurally unmatched and matched stimuli on immediate and subsequent vocal stereotypy that was displayed by three children with autism spectrum disorders. For 2 of the 3 participants, access to matched stimuli, unmatched stimuli, and music decreased immediate levels of vocal stereotypy; however, with the exception of matched stimuli for one participant, none of the stimuli produced a clear abolishing operation for subsequent vocal stereotypy. That is, vocal stereotypy typically increased to baseline levels shortly after alternative stimulation was removed. Detection of motivating operations for each participant's vocal stereotypy was aided by the analysis of component distributions. The results are discussed in terms of immediate and subsequent effects of preferred stimuli on automatically reinforced problem behavior.

Keywords: *automatic reinforcement, motivating operations, multiple-schedule, noncontingent stimulation, stereotypy*

Authors' Note: This study was completed in partial fulfillment of a Masters degree in Behavior Analysis by the first and second authors at Saint Cloud State University, St. Cloud, MN. We thank Catherine Gale, Diane Hayward, Dr. Svein Eikeseth, and Dean Smith from the UK Young Autism Project for their support and guidance throughout. Address correspondence to John T. Rapp, Education Building, A261, St. Cloud State University, 720 Fourth Avenue South, St. Cloud, MN 56301-4498; email: jtrapp@stcloudstate.edu

Although several recent studies on automatically reinforced behavior have included novel assessments or interventions (Ahearn, Clark, MacDonald, & Chung, 2007; Rapp, 2007; Roantree & Kennedy, 2006), research on nonsocially reinforced behavior clearly lags behind research on socially reinforced behavior. One reason for this discrepancy appears to be linked to challenges posed by controlling the consequence maintaining automatically reinforced behavior. Because such behavior generates its own reinforcing stimulation, experimenters must block or prevent the behavior to control its consequence. Unless protective equipment can be used to attenuate the sensory consequence(s) maintaining the behavior (Iwata, Pace, Cowdery, & Miltenberger, 1994; Rapp, Miltenberger, Galensky, Ellingson, & Long, 1999), reinforcement for such behavior remains available.

Several studies have shown that baseline measures of automatically reinforced behavior vary considerably on a day-to-day basis (Cuvo, May, & Post, 2001; Rapp, Vollmer, Dozier, St. Peter, & Cotnoir, 2004; Vollmer, Marcus, & LeBlanc, 1994). Variability in day-to-day measures of stereotypy may pose severe challenges for researchers who assess and treat automatically reinforced behavior in at least two ways. First, high variability may obscure or mask patterns produced by the introduction of an intervention that does not reduce the behavior to zero. Second, high variability may cause considerable overlap across phases, which weakens the demonstration of control over the behavior. One potential solution to this problem is to evaluate automatically reinforced behavior using a relative measure. The three-component multiple-schedule described by Simmons, Smith, and Kliethermes (2003) compares level (i.e., frequency or duration) of a behavior across three consecutive components. Typically, the first and third components are baseline conditions, and the second component is either a baseline or treatment condition. The level of the behavior in one component can be compared to the level of the behavior in the other two components, and the components with the lowest and highest levels of the behavior can be identified. This measure may be described as "relative" because the rank-order of a component (e.g., lowest or highest) is dependent on the level of the behavior in the other two components. Relative measures of the behavior (i.e., distribution of the behavior across components) may be more reliable and sensitive than absolute measures (i.e., frequency or duration alone) because the variables that may be responsible for day-to-day variability should affect all three components of the multiple-schedule.

For example, assume that an individual exhibits stereotypy for 35%, 80%, 25%, and 50% of a 30-min observation period across 4 consecutive days. In this case, visual analysis that is based on line graphs would indicate that stereotypy was highly variable. Moreover, visual analysis would suggest that additional baseline sessions would be needed before the effects of an independent variable could be adequately evaluated. By contrast, analysis of this same data using the three-component method may indicate that the rank-order of the components was relatively stable across days (e.g., the first component was consistently the highest) and additional sessions were not necessary. Stability in rank-order would allow the introduction of an independent variable in one or multiple components. When used in conjunction with a brief reversal design, shifts in the distribution across components (i.e., changes in the rank-order of components) may be attributed to the effects of an independent variable.

A few studies have used three-component multiple-schedules to assess the effects of motivating operations (MOs) on automatically reinforced behavior (Cuvo et al., 2001; Rapp, 2006, 2007; Simmons et al., 2003). Motivating operations are events that alter the value of a consequence and the occurrence of behavior that has been associated with that consequence (Laraway, Snycerski, Michael, & Poling, 2003). Laraway et al. described two types of MOs: the abolishing operation (AO), which decreases the value of a consequence and the establishing operation (EO), which increases the value of a consequence. When evaluating automatically reinforced behavior, presence of an EO is inferred by high or increased levels of the behavior and the presence of an AO is inferred by low or decreased levels of the behavior (Rapp, 2007). Redistribution of automatically reinforced behavior across the three components of a multiple-schedule allows the experimenter to identify stimuli that exert immediate (i.e., when a preferred stimulus is present) and subsequent (i.e., after a preferred stimulus is removed) effects on automatically reinforced behavior. To this end, the three-component multiple-schedule may serve as a sensitive tool for assessing the operant function of problem behavior.

In recent years, several studies have examined the effects of antecedent-based interventions on automatically reinforced behavior (Ahearn, Clark, DeBar, & Florentino, 2005; Britton, Carr, Landaburu, & Romick, 2002; Higbee, Chang, & Endicott, 2005; Piazza, Adelinis, Hanley, Goh, & Delia, 2000; Rapp, 2006; Rapp et al., 2004). Antecedent procedures often involve the use of noncontingent reinforcement (NCR), which consists of providing noncontingent access to stimuli that are identified via empirical preference stimulus assessments (LeBlanc, Patel, & Carr, 2000). In terms of treating automatically reinforced behavior, NCR can be divided into two broad intervention categories. The first category, structurally unmatched stimulation, provides noncontingent access to preferred, but arbitrary stimuli that are not necessarily matched to the putative sensory product of stereotypy. The second category, structurally matched stimulation, consists of stimuli that are intended to match the putative sensory product of stereotypy.

Based on the results of empirical preference assessments, Piazza et al. (2000) found that structurally matched stimuli provided greater decreases in automatically reinforced behavior than structurally unmatched stimuli; however, Ahearn et al. (2005) found the converse. Nevertheless, results of both studies are potentially limited insofar as neither of the studies assessed the subsequent levels of stereotypy (i.e., after the matched or unmatched stimulus was removed) to determine if preferred stimuli were functionally matched to the problem behavior. For example, it is possible that a structurally matched stimulus does not produce a sensory consequence that is functionally matched to the automatically reinforced target behavior. Thus, automatically reinforced behavior may increase above preintervention levels after the preferred stimulus is removed (Rapp, 2006, 2007). Analysis of automatically reinforced behavior using the three-component multiple-schedule may elucidate possible changes in MOs as a function of preferred stimulation and help clarify the different findings from the Piazza et al. and Ahearn et al. studies.

Treatment for automatically reinforced behavior should yield two outcomes. First, problem behavior should decrease when the preferred stimulus is present. Second, removal of the preferred stimulus should be marked by either (a) the continued reduction or further decreases in the problem behavior or (b) an increase in problem behavior that does not exceed the level that is expected if access to preferred stimuli had not been provided. From a clinical standpoint, the pattern described in (a) is most desirable as the effects of the intervention are shown to endure after the preferred stimulus is removed; however, the outcome described in (b) is also desirable insofar as the overall amount of time the individual engaged in problem behavior is reduced. Conceptually, both patterns indicate that preferred stimulation produced an AO for the problem behavior. Conversely, an increase in problem behavior that exceeds the level that is expected if access to preferred stimuli had not been provided would indicate that such stimulation produced an EO for problem behavior (i.e., the preferred stimulus was not functionally matched to the product of the behavior); this pattern is not clinically desirable.

Because identifying AOs for automatically reinforced behavior may lead to more effective treatment procedures, a methodology for identifying such effects should be developed. Likewise, as the results from some recent studies suggested that behavior changes that are produced by some events may be relatively small (Rapp, 2006, 2007), it may be reasonable to use relative levels of stereotypy as a supplemental dependent measure to help facilitate the identification of potentially important variables. The purposes of the present study were to (a) identify structurally unmatched and matched preferred items that may compete with engagement in stereotypy, (b) determine if unmatched and matched stimuli decrease immediate stereotypy and function as AOs or EOs for subsequent stereotypy, and (c) use a methodology that focuses on the relative changes in stereotypy across the three components of a multiple-schedule to supplement visual analysis of data. Based on the results obtained by Rapp (2006, 2007), we expected the structurally matched stimuli to be functionally matched to the sensory product maintaining the behavior and, thereby, to produce AOs for subsequent engagement in vocal stereotypy. In contrast, we expected the unmatched stimuli to compete with the sensory product maintaining the behavior and function as EOs for subsequent engagement.

Method

Participants and Settings

Three children who were diagnosed with autism spectrum disorders participated in the experiment. Dave, Adam, and Bobby were diagnosed by their pediatrician using the ICD-10 criteria at the ages of 26, 29, and 25 months, respectively. Bobby was a 7-year-old boy who used three- to five-word utterances to request items and one- to two-word utterances to label objects. Dave was a 7-year-old boy who had no expressive language and inconsistently followed simple one-step instructions. Adam was a 5-year-old boy who used four- to five-word utterances to label and request objects and followed two- to three-step instructions. Each participant displayed vocal stereotypy. The sessions were conducted three to five times per week at various times during the day (between the hours of 9 a.m. and 5 p.m.) in a workroom in each child's home. One 30-min session, which contained three 10-min components, was conducted per day with each participant. Apart from the furniture already present in the workrooms and the stimuli described below, no additional materials were available during the sessions.

Data Collection, Response Definitions, and Reliability

Data were collected on the duration of vocal stereotypy and of toy manipulation. *Vocal stereotypy* was defined as acontextual audible sounds or words produced with an open or closed mouth. *Toy manipulation* was defined as the individual touching an object with any part of his hand. For toys that produced music, the offset of toy manipulation was recorded when the toy stopped playing music (after approximately 10 s). All sessions were videotaped and subsequently scored by trained graduate students using laptop computers equipped with data collection programs.

A second observer scored at least 40% of the sessions for each participant. Interobserver agreement (IOA) was calculated using the proportional method by dividing each session into 10-s bins. Agreement scores were calculated for each bin by dividing the lower duration by the higher duration and then multiplying that value by 100%. The mean IOA score for each dependent event was calculated by dividing the sum of all the bins by the total number of bins. For each participant, the IOA score for toy manipulation was produced by averaging the scores from across all toys during stimulus preference assessment sessions and the second component of the respective intervention sequences. For Bobby, mean IOA scores were 97.2% (range, 95.9% to 98.3%) for vocal stereotypy and 99.2% (range, 98.7% to 99.9%) for toy manipulation. For Dave, the mean IOA scores were 96.7% (range, 93.4% to 98.8%) for vocal stereotypy and 99.6% (range, 99.3% to 100%) for toy manipulation. For Adam, mean IOA scores were 92.6% (range, 88.5% to 99.6%) for vocal stereotypy and 99.4% (range, 99.1% to 99.6%) for toy manipulation.

Procedures and Experimental Design

Prior to the assessment of vocal stereotypy, a free-operant stimulus preference assessment (FOSPA) was conducted to identify highly preferred stimuli for each participant (Roane, Vollmer, Ringdahl, & Marcus, 1998). Subsequently, effects of preferred stimuli on each participant's vocal stereotypy were evaluated using a combination of a three-component multiple-schedule and a brief reversal design. We manipulated the events in the second component of the three-component multiple-schedule to determine the effects on immediate (i.e., within the second component) and subsequent (i.e., within the third component) levels of vocal stereotypy.

Stimulus preference assessment. Three or more FOSPA sessions were conducted with each participant to identify preferred stimuli that may compete with vocal stereotypy. Sessions were 15 min in duration. Each child was placed in a room with 8 to 10 stimuli on the floor. Prior to beginning the preference assessment, the observer prompted the child to sample each stimulus for 10 s. No social consequences were provided during these sessions. The stimuli were selected based on parental interviews, availability of toys, and on the putative stimulation that was generated by vocal stereotypy.

The stimuli evaluated for Bobby were a toy camera, books, a toy phone, a keyboard, a singing hamster, an owl toy, a robin toy, a robot, a Bob the Builder television toy, and a ring stacker. Stimuli evaluated for Dave were bells, a squidgy ball, a rainmaker, a glitter stick, a spiky ball, a wand, a helicopter, a rattle, a hand massager, and maracas. The stimuli evaluated for Adam were play-doh, a fifi car, a magazine, a flashing ball, a la la cuddly toy, a rattle, a caterpillar ring stacker, a face bat, a glitter tube, and a fimble toy. Three sessions were initially conducted for each participant. Five sessions were conducted with Adam because he allocated most of his responding to one or two toys during the initial FOSPA sessions. Starting at the third session, specific toys were removed from the FOSPA to increase the probability that Adam would interact with toys that matched the putative sensory product of vocal stereotypy. Specifically, the flashing ball was removed starting session 3, the la la cuddly toy starting session 4, and the play-doh for session 5.

For Adam, we conducted a reinforcer assessment because his initially low levels of interaction with musical toys suggested that auditory stimuli may not function as reinforcers. The combination of a concurrent operants and a brief ABA reversal design was used to determine whether music functioned as a reinforcer for an arbitrary behavior (Fisher et al., 1992). The reinforcer assessment lasted 15 min and was divided into 3, 5-min phases. We positioned two identical chairs at the extreme ends of a room and recorded the duration of time Adam sat in each chair. During min 1 through 5, no consequences were provided for sitting in either chair. During min 6 through 10, music was provided contingent on sitting in one of the two chairs (i.e., the music chair) and no consequences were delivered for sitting in the other chair (i.e., the control chair). Finally, during min 11 through 15, no consequences were provided for sitting in either chair. If music was a reinforcing stimulus for Adam's behavior, then the percentage of time that he sits in the designated chair should increase when music is provided contingent on sitting in that chair.

Assessment of MOs. The effects of structurally unmatched and matched stimulation on each participant's stereotypy were evaluated using three-component multiple-schedules combined with a brief reversal design (with four sequences) wherein two sequences were alternated in a pairwise fashion. Sessions were 30 min in duration for each participant (3, 10-min components). In regard to the components in the multiple-schedule, the free-operant (FO) components were signaled by the absence of alternative stimulation whereas the unmatched toy, matched toy, and music components were signaled by the presence of the respective stimulus. In all sequences, the first and third components were FO components during which each participant was placed in a room with no preferred stimuli and no social consequences were provided. In the FO sequence, the second component was also a FO component. In the unmatched (UN) sequence, an empirically identified preferred stimulus that did not match the putative sensory product of stereotypy was presented during the second component of the multiple-schedule. In the matched (MA) sequence, an empirically identified preferred stimulus that matched the putative sensory product of stereotypy was presented during the second component of the multiple-schedule.

In addition, because the results obtained by Rapp (2007) suggested that music was a functionally matched stimulus for the vocal stereotypy of two participants, we assessed the effects of providing music in order to evaluate another structurally matched stimulus for each participant's vocal stereotypy. Moreover, because of the relative ease of identifying music as a structurally matched stimulus for vocal stereotypy, it appeared to be an ideal stimulus with which to further evaluate the three-component methodology. In the music (MU) sequence, noncontingent access to music was provided during the second component of the multiple-schedule.

Both unmatched and matched stimuli were selected based on the results from FOSPAs. The structurally unmatched stimuli were books for Bobby, the squidgy ball for Dave, and the flashing ball for Adam. The structurally matched stimuli were the keyboard for Bobby, the bells and the rattle for Dave, and the fifi car and the rattle for Adam. For each participant, the matched stimuli were selected because each generated auditory stimulation whereas the unmatched stimuli were selected because none generated auditory stimulation. Each participant was exposed to five MA sequences, five UN sequences, two or three MU sequences, and six or more FO sequences. Data from one FO sequence for Dave (which was conducted between sessions 15 and 16) were omitted because vocal stereotypy was below 3% of the session time across all three components.

Data Analysis

The effects of various stimuli on each participant's stereotypy were evaluated via visual inspection of absolute and relative levels of vocal stereotypy in each component within and across sequences. In the FO sequence, the participant did not have access to preferred stimulation during any of the three components of the multiple-schedule. In the Intervention sequences (i.e., Unmatched, Matched, and Music), the child had access to preferred stimulation during the second component of the multiple-schedule, and the first and third components remained procedurally the same as in the FO sequence.

We used line graphs to examine whether absolute levels of vocal stereotypy were differentiated across the second and third components for each pairwise comparison. If levels of vocal stereotypy were undifferentiated or if the difference between the two sequences was small in the second component, we examined relative levels of vocal stereotypy between-sequences using a component distribution (i.e., a graph of the proportion of sessions wherein levels of stereotypy were lowest and highest in each component and higher in the first component than in the third component). That is, the relative levels of vocal stereotypy between the second components from different sequences were analyzed on a session-by-session basis. If results were undifferentiated during the third component, we conducted a combined within-sequence and between-sequence analysis by comparing preintervention with postintervention levels of vocal stereotypy across sequences.

Line graphs. Data on the percentage of time each participant engaged in vocal stereotypy during the second component were compared across sequences. This analysis was repeated for the third component of the respective sequences.

Between-sequence analysis. The between-sequence analysis involves comparisons of the patterns from two sequences in order to make conclusions about the effects of an independent variable (i.e., preferred stimuli). The relative levels of stereotypy during the Intervention and FO sequences were compared to determine whether different types of stimulation decreased immediate levels of stereotypy. Specifically, the proportion of sessions that the second component was the lowest and the highest was compared across sequences. For example, a pattern wherein the second component was more often the lowest and less often the highest in an Intervention sequence than in a FO sequence would indicate that preferred stimulation functioned as an AO for immediate vocal stereotypy. Conversely, a pattern wherein the second component was more often the highest and less often the lowest in an Intervention sequence than in a FO sequence would indicate that preferred stimulation functioned as an EO for immediate vocal stereotypy.

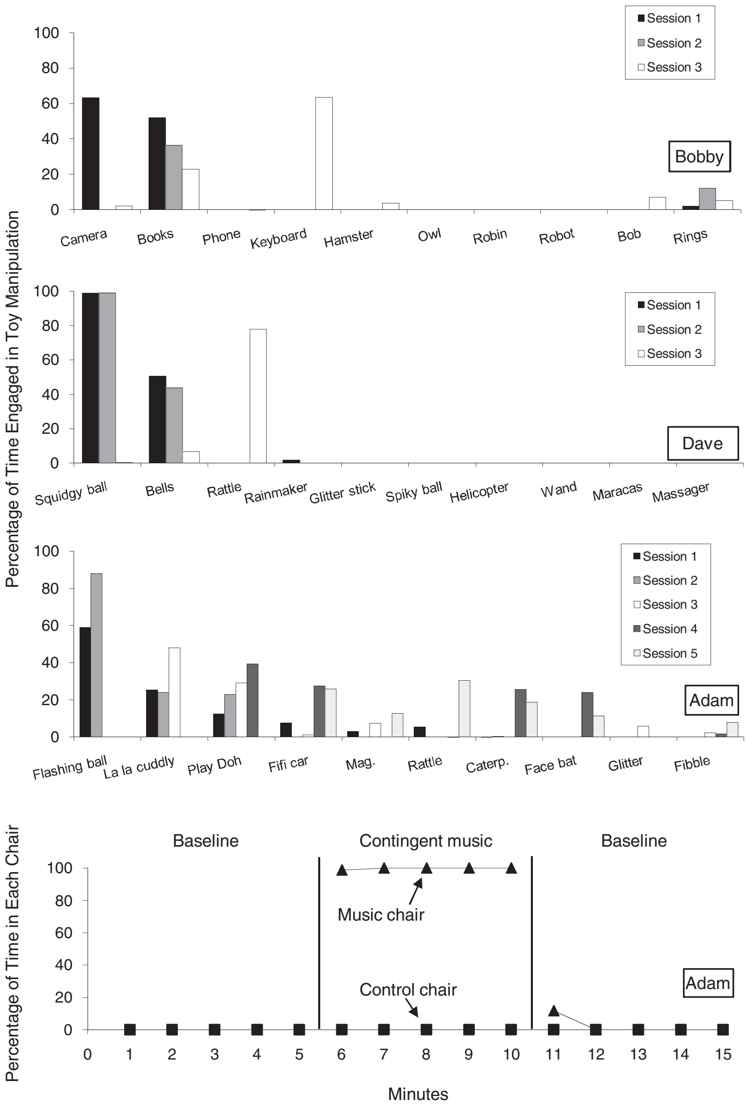
Combination of within- and between-sequence analyses. First, a within-sequence analysis was used to examine patterns that were observed across components within FO sequences and Intervention sequences. The proportion of sessions the first component was higher than the third component was computed for each sequence. In this way, the FO and Intervention sequences showed whether stereotypy decreased or increased across components. Next, we conducted a between-sequence analysis by comparing relative levels in each sequence. For example, a pattern wherein the first component was higher than the third component more often during the Intervention sequence than during the FO sequence would suggest that preferred stimulation produced an AO for subsequent vocal stereotypy in the third component (i.e., vocal stereotypy decreased or remained low following the removal of preferred stimulation). Conversely, a pattern wherein the first component was higher than the third component more often during the FO sequence than during the Intervention sequence would suggest that preferred stimulation produced as an EO for vocal stereotypy in the third component (i.e., vocal stereotypy increased following the removal of preferred stimulation). Finally, a pattern wherein the distribution of stereotypy across the first and third components of the Intervention sequence was similar to the distribution of vocal stereotypy across the components of the FO sequence would suggest either (a) a mild AO effect (i.e., access to the preferred stimulus was comparable to having access to stereotypy) or (b) no MO effect.

Results

Figure 1 depicts the results from the FOSPA for Bobby (first panel), Dave (second panel), and Adam (third panel). Bobby's preferred unmatched stimuli were the books (*M* = 37%) and his preferred matched stimulus was the keyboard (*M* = 21%). Figure 1 (second panel) shows that Dave manipulated the squidgy ball (*M* = 66%) the most among the unmatched stimuli and the bells (*M* = 50%) and the rattle (*M* = 26%) the most among the matched stimuli. Adam manipulated the flashing ball (*M* = 73%) the most during the first two sessions. Thereafter, unmatched stimuli were removed until a matched stimulus, the rattle (in session 5), which was manipulated for 30% of the last session, became the most preferred.

Figure 1 (fourth panel) shows results of the reinforcer assessment for Adam. During the baseline period (min 1-5), Adam did not sit in either chair. During the contingent music period (min 6-10), the percentage of

Figure 1
Percentage of time engaged in object manipulation during stimulus preference assessments for Bobby (first panel), Dave (second panel), and Adam (third panel), and percentage of time sitting in each chair during the reinforcer assessment of music for Adam (fourth panel)

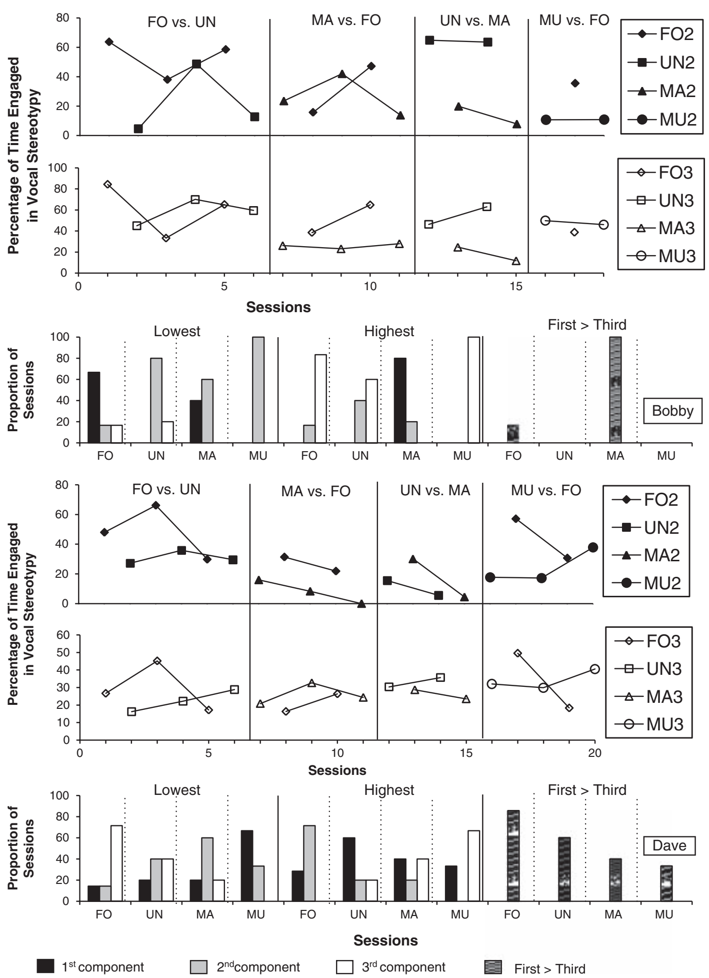


time that Adam sat in the "music" chair increased from zero to nearly 100%. During the return to baseline (min 11-15), sitting in the music chair decreased to near-zero levels. Although this assessment was brief, the results show that music functioned as a reinforcer for an arbitrary behavior (i.e., sitting in a chair) and suggest that music was a preferred stimulus for Adam.

Figure 2 depicts the line graphs with the second (closed data points) and third components (opened data points) and the component distributions for Bobby's (upper three panels) and Dave's (lower three panels) engagement in vocal stereotypy during FO, UN, MA, and MU sequences. For Bobby, Figure 2 (first panel) shows that during the second component, levels of vocal stereotypy were generally higher in the FO sequence than in the UN sequence during the FO versus UN comparison, suggesting that the unmatched stimuli decreased vocal stereotypy. Although the data paths for the second components in the MA versus FO sequences were undifferentiated, the UN versus MA comparison shows that the matched stimulus produced lower levels of vocal stereotypy than the unmatched stimuli. Visual analysis of the data across comparisons (sessions 2 through 14) indicates that the unmatched stimuli became less effective for decreasing stereotypy across sessions. The data path for the second component of the MU sequence was clearly lower than the FO data point, suggesting that music also decreased vocal stereotypy. Between-sequence analyses of the component distributions, which were used to clarify the results from the FO versus UN and MA versus FO comparisons, show that the second component was lowest for 1 of 6 sessions (17%) in the FO sequence compared to 4 of 5 sessions (80%) in the UN sequence and 3 of 5 sessions (60%) in the MA sequence. These patterns suggest that both unmatched and matched stimuli decreased immediate vocal stereotypy.

Figure 2 (second panel) shows that the FO versus UN comparison did not yield differentiated data paths for the third component for Bobby. By contrast, the third component of the MA sequence was always lower than the third component of the FO sequence during the MA versus FO comparison, suggesting that the matched stimulus functioned as an AO for subsequent stereotypy. Likewise, the third component of the MA sequence was consistently lower than the third component of the UN sequence. The MU versus FO comparison was difficult to interpret because levels of vocal stereotypy were only marginally higher in the MU sequence than in the FO session. Within- and between-sequence analyses (third panel) of the component distribution show that the first component was higher than the third component for 1 of 6 sessions (17%) during the FO sequence, 0 of 5

Figure 2
The percentage of time engaged in vocal stereotypy during the second (closed data points) and the third (opened data points) components of FO, Unmatched (UN), Matched (MA), and Music (MU) sequences for Bobby (first and second panels) and Dave (fourth and fifth panels). Component distributions of the proportion of sessions for which vocal stereotypy was lowest and highest in each component, and higher in the first component than in the third component for Bobby (third panel) and Dave (sixth panel)



sessions (0%) during the UN sequence, 5 of 5 sessions (100%) during the MA sequence, and 0 of 2 sessions (0%) during the MU sequence. As a whole, these results suggest that unmatched stimuli and music produced EOs for subsequent stereotypy and structurally matched stimulation produced an AO for subsequent vocal stereotypy. Thus, only the structurally matched stimulus was functionally matched to Bobby's vocal stereotypy.

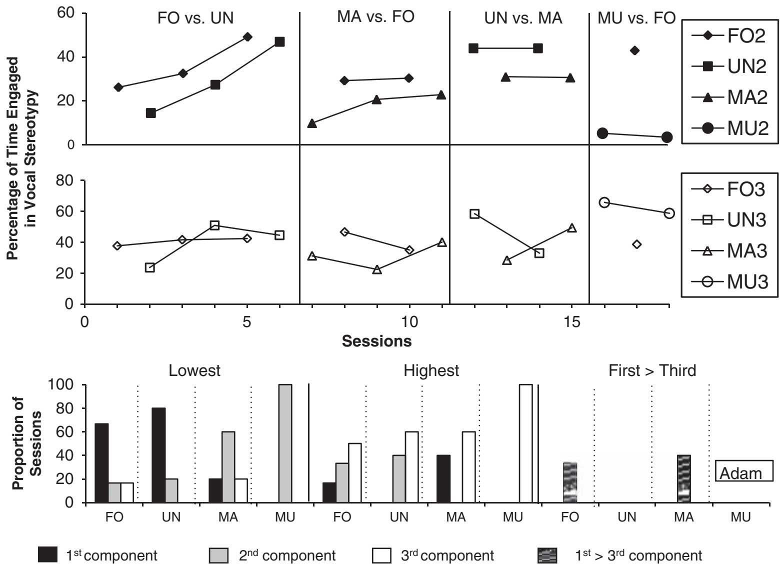
For Dave, visual analysis of Figure 2 (fourth panel) shows that levels of vocal stereotypy were generally higher in the second component of the FO sequence than in second component of the UN and MA sequences, suggesting that both unmatched and matched stimuli decreased stereotypy in the second component. Results of the UN versus MA comparison were difficult to interpret because only two sessions were conducted with each sequence and the data paths converged at near-zero levels; however, the results suggest that both unmatched and matched stimuli produced low levels of immediate vocal stereotypy for Dave. In the final comparison of the second components, Dave's vocal stereotypy was generally lower in the MU sequence than in the FO sequence, but the decreasing trend in the FO sequence made this comparison difficult to interpret. However, between-sequence analyses of the component distributions show that the second component was lowest for 1 of 7 sessions (14%) during the FO sequence, 2 of 5 sessions (40%) during the UN sequence, 3 of 5 sessions (60%) during the MA sequence, and 1 of 3 sessions (33%) during the MU sequence. Conversely, the second component was highest for 5 of 7 sessions (71%) during the FO sequence, 1 of 5 sessions (20%) during the UN sequence, 1 of 5 sessions (20%) during the MA sequence, and 0 of 3 sequences (0%) during the MU sequence. These patterns indicate that unmatched stimuli, matched stimuli, and music decreased Dave's immediate engagement vocal stereotypy. However, these patterns did not clarify whether either matched or unmatched stimulation was more effective at decreasing immediate stereotypy for Dave.

Figure 2 (fifth panel) shows that the data paths for Dave's vocal stereotypy in the third components of the FO versus UN, MA versus FO, and MU versus FO comparisons were undifferentiated, which suggests that the effects of matched stimuli, unmatched stimuli, and music were comparable to those produced by prior access to stereotypy. Although the UN versus MA comparison contained only two sessions with each sequence, data from the prior comparisons indicated that UN data path was on an increasing trend across sessions whereas the MA data path was relatively stable. Within- and between-sequence analyses of the component distribution (sixth panel) show that the first component was higher than the third component for 6 of 7 sessions (86%) during the FO sequence, 3 of 5 sessions (60%) during the UN sequence, 2 of 5 sessions (40%) during the MA sequence, and 1 of 3 sessions (33%) during the MU sequence. These results suggest that music produced an EO for subsequent vocal stereotypy and, therefore, it was not a functional match to vocal stereotypy. The effects of the unmatched and matched stimuli on subsequent stereotypy were less clear because vocal stereotypy was typically lower in the third component than in the first component of the FO sequence. Nevertheless, when the effects of unmatched and matched stimuli are compared to the effects of prior access to stereotypy, the unmatched and matched stimuli should be considered EOs for subsequent vocal stereotypy because each increased subsequent engagement in vocal stereotypy relative to the FO sequence.

Figure 3 shows the component distribution for Adam and the percentage of time he engaged in vocal stereotypy across FO, UN, MA, and MU sequences. With the possible exception of the increasing trends in the FO versus UN comparison, data from the second component (first panel) show that vocal stereotypy was higher in the FO sequence than in each intervention sequence, suggesting that unmatched stimuli, matched stimuli, and music decreased Adam's vocal stereotypy. Furthermore, the differentiation between the data paths of the second components during the UN versus MA comparison suggests that the matched stimuli were more effective than the unmatched stimulus in decreasing stereotypy. Because the differences between the sequences were relatively small during the FO versus UN and MA versus FO comparisons, we conducted between-sequence analyses using the component distributions. The second component was lowest for 1 of 6 sessions (17%) during the FO sequence, 1 of 5 sessions (20%) during the UN sequence, and 3 of 5 sessions (60%) during the MA sequence. Conversely, the analyses show that vocal stereotypy was highest in the second component for 2 of 6 sessions (33%) during the FO sequence, 2 of 5 sessions (40%) during the UN sequence, and 0 of 5 sessions (0%) during the MA sequence. These patterns suggest that the matched stimuli decreased immediate vocal stereotypy and the unmatched stimulus did not consistently decrease immediate vocal stereotypy.

Figure 3 (second panel) shows that during the third component for Adam's vocal stereotypy, the data paths in the FO versus UN comparison remained undifferentiated. Furthermore, the third component of the MA sequence was generally lower than the FO sequence, but the decreasing trend in the FO sequence in the MA versus FO comparison and the undifferentiated pattern in the UN versus MA comparison suggested that further

Figure 3
Percentage of time engaged in vocal stereotypy during the second (closed data points) and the third (opened data points) components of FO, Unmatched (UN), Matched (MA), and Music (MU) sequences for Adam (first and second panels). Component distribution of the proportion of sessions for which vocal stereotypy was lowest and highest in each component, and higher in the first component than in the third component for Adam (third panel).



analysis with the component distributions was warranted. In the final comparison, vocal stereotypy during the third component was clearly higher in the MU sequence than in the FO sequence, suggesting that music functioned as an EO for subsequent stereotypy. That is, music was not functionally matched to the consequence of Adam's vocal stereotypy. A combination of within- and between-sequence comparisons (third panel) shows that the first component was higher than the third component for 2 of 6 sessions (33%) during the FO sequence, 0 of 5 sessions (0%) during the UN sequence, and 2 of 5 sessions (40%) during the MA sequence. The subsequent effects of structurally unmatched and matched stimuli were not clear from these analyses because Adam's vocal stereotypy typically increased across components. At a minimum, the results suggest that structurally unmatched and matched stimuli did not produce AOs for subsequent vocal stereotypy. Thus, these stimuli were not functionally matched to the product of Adam's vocal stereotypy.

Discussion

As a whole, results from the brief experimental analyses showed that (a) structurally matched stimulation and music decreased immediate vocal stereotypy for all three participants, (b) structurally matched stimuli produced larger decreases in vocal stereotypy than unmatched stimuli for two of three participants, (c) structurally unmatched stimuli decreased immediate vocal stereotypy (at least temporarily) for two of three participants, and (d) despite the finding that unmatched stimuli were more preferred, unmatched stimuli did not decrease subsequent stereotypy for any of the three participants. In fact, matched stimuli produced lower levels of vocal stereotypy than unmatched stimuli in the second and third components for two of three participants (the levels of stereotypy for the other participant were comparable in both components of both sequences).

Structurally matched stimulation decreased immediate vocal stereotypy for all three participants, but only decreased subsequent stereotypy for Bobby. In this way, matched stimuli competed with but did not substitute for the stimulation generated by vocal stereotypy for the other two participants. In a similar vein, results for music with each participant provide clear support for the need to conduct empirical assessments of functionally matched stimulation. In general, results indicate that labeling a stimulus as structurally "unmatched" (based on the dissimilar stimulation it produces) correctly predicted that it was not functionally matched to the product of stereotypy; however, the converse was not true for stimuli that were labeled as structurally "matched."

The combined use of line graphs and component distributions to examine within- and between-sequence patterns facilitated a more comprehensive, conceptual analysis of the results. That is, line graphs of the second and third components highlighted trends in data across sessions, but did not allow the comparison of patterns within-sequences (e.g., identifying whether prior access to stereotypy functioned as an AO for subsequent stereotypy). In contrast, component distributions facilitated the detection of within-sequence patterns, but did not detect trends across sessions. Thus, data generated by component distributions and line graphs, in combination, lead to the detection of changes in MOs for vocal stereotypy in the presence or following the removal of alternative stimuli for each participant.

Interestingly, all three participants showed stronger preferences for structurally unmatched stimuli than for structurally matched stimuli. Ahearn et al. (2005) suggested that because stimulation from automatically reinforced behavior was always available, individuals who display automatically reinforced problem behavior may prefer stimulation that differs from the stimulation produced by their behavior. Considering that each participant exhibited relative preferences for unmatched stimuli, showing that structural similarity does not necessarily predict functional similarity has important clinical implications. Specifically, these results suggest that selecting stimuli based on an individual's preference and whether the stimulus competes with (i.e., decreases) stereotypy is not sufficient for ensuring that the removal of the stimulus will not produce an EO for subsequent stereotypy. To this end, the three-component multiple-schedule appears to be a potentially useful assessment tool for differentiating between structurally and functionally matched stimuli.

Although music decreased vocal stereotypy for each participant, it was not an AO for subsequent stereotypy for Bobby, Dave, or Adam. Our results differ from those obtained in the Rapp (2007) study, which showed that music was functionally matched to the stimulation generated by vocal stereotypy. It is possible that vocal stereotypy displayed by the participants in the current study was reinforced by specific frequencies or qualities of sound that were not produced by music or toys. Given that music produced an EO for Bobby's subsequent vocal stereotypy, it is not clear why the matched stimulus, which also produced auditory stimulation, functioned as an AO for Bobby's subsequent vocal stereotypy (see Figure 2, second panel). It is possible that the structurally matched toy (a keyboard) allowed Bobby more control over the auditory stimulation that was generated by the toy (see Rapp, 2008 for a discussion of conjugate reinforcement). That is, because Bobby controlled the auditory stimulation generated by the keyboard, the stimulation may have been more similar to the product of his vocal stereotypy.

Results from this study extend the literature on the use of matched and unmatched stimuli to treat automatically reinforced behavior in at least three ways. First, this study found that individuals tended to prefer (based on behavior allocation across three or more sessions) stimuli that did not match the putative sensory product of their stereotypy. Ahearn et al. (2005) found that unmatched stimuli were more effective for decreasing automatically reinforced behavior than matched stimuli; however, Ahearn et al. did not evaluate relative preferences for unmatched and matched stimuli. Matched stimuli that were used in the Ahearn et al. study may have been considerably less preferred by the participants than the unmatched stimuli, which would explain why their results were different. Second, the results of this study showed that structurally matched stimuli were not necessarily functionally matched to the product of vocal stereotypy. Specifically, although music competed with vocal stereotypy for two of the three participants, it also produced a clear EO for subsequent vocal stereotypy by Adam. Third, this study extends the methodology that was described by Simmons et al. (2003) by employing relative measures of behavior within a three-component multiple schedule to minimize the influence of day-to-day variability when assessing the effects of interventions on automatically reinforced behavior.

The within-sequence patterns observed in this study were inconsistent with the results from the Rapp (2004, 2007) studies insofar as prior access to stereotypy did not function as an AO for subsequent stereotypy for two of the three participants. Although Ahearn, Clark, Gardenier, Chung, and Dube (2003) suggested that stereotypy gains “momentum” across time, it is possible that the duration of time required to produce decreasing levels of stereotypy across components varies across individuals. Given that most stereotypy appears to be maintained by automatic positive reinforcement (see Rapp & Vollmer, 2005), a decrease in stereotypy across some unit of time is expected. The use of components with longer durations may have allowed such patterns to emerge.

For the most part, the behavior changes observed in this study were relatively small. The study of small behavior changes has not been a hallmark of applied behavior analysis (Baer, 1977; Baer, Wolf, & Risley, 1968); however, variables that initially produce small behavior changes may serve as catalysts for larger, subsequent changes. For example, results from the third component of the MA versus FO comparison for Bobby (see Figure 2, second panel) show that he did *not* display vocal stereotypy for approximately 7.5 min of each component after the matched stimuli were available compared to approximately 5 min of each component after matched stimuli were not available. Thus, recent access to the matched stimuli produced an additional 2.5 min window of time per session during which the value of stimulation produced by Bobby’s vocal stereotypy was lower and, thereby, additional trials of academic training could have been conducted. Assuming that the increased number of academic trials led to the acquisition of new behaviors, such behavior may slowly displace and ultimately compete with stereotypy for behavior allocation. Likewise, packaging procedures that exert small behavior changes together (e.g., prior access to functionally matched stimuli and contingent verbal reprimands) may lead to socially significant changes in behavior.

Some potential limitations to the results of the current investigation should also be discussed. First, the component distributions involved data that were collapsed from several sessions, which potentially obscured some behavior patterns (e.g., increasing or decreasing trends). As such, we recommend that component distributions be used in combination with other graphs that depict trends. Second, the order in which the sequences were conducted was not counterbalanced across participants (i.e., the UN sequence was always introduced before the MA sequence) because we exposed each participant to his most preferred stimuli, which were unmatched, before exposing each to less preferred stimuli. Similarly, only two sessions were conducted with each sequence for some comparisons (e.g., UN vs. MA). Nevertheless, the same number of sessions was conducted with the MA and UN sequences and the effects of matched and unmatched stimuli were directly compared in a pairwise fashion. Thus, it is unlikely that the order of the sequences influenced the results of this study.

Third, because the participants' preferences for music were not evaluated with an empirical preference assessment, it is not clear if music was more or less preferred than the structurally unmatched items. Nonetheless, access to music decreased vocal stereotypy for two of the three participants and was shown to reinforce an arbitrary behavior for Adam. Finally, we did not conduct a traditional functional analysis (Iwata, Dorsey, Slifer, Bauman, & Richman, l994) to demonstrate that each participant's problem behavior was automatically reinforced. Nevertheless, Vollmer, Marcus, Ringdahl, and Roane (1995) proposed a methodology wherein consecutive no-interaction sessions were used to demonstrate the persistence of problem behavior after the results of a functional analysis were undifferentiated. Using the same logic, we simply conducted no-interaction or FO sessions as a control sequence (against which the other sequences were compared) to demonstrate the persistence of each participant's vocal stereotypy in the absence of social consequences.

Results from this study provide several avenues for future research for evaluating problem behavior using relative measures. First, future studies could determine the optimal component duration that is necessary to identify relative changes in the value of automatically reinforced behavior across time. Second, future research should use the three-component multiple-schedule to evaluate the effects of other procedures on subsequent levels of stereotypy. For example, Ahearn et al. (2007) used a response interruption and redirection (RIRD) procedure to decrease vocal stereotypy; however, the effects of the procedure on subsequent stereotypy were not evaluated. Thus, the extent to which the removal of the RIRD intervention produced EOs or AOs for subsequent vocal stereotypy is not known. Third, to determine the generality of the effects described in this study, researchers should conduct direct and systematic replications of the procedures with socially and nonsocially reinforced behavior. To improve on the methodology used in this study, we recommend that future researchers focus on obtaining stability in the second component and then observing for potential behavior changes (i.e., EOs or AOs) in the third component. To date, most research has focused on structurally matched and unmatched stimuli because functionally matched stimuli were difficult to identify using traditional single-subject methodology. However, the use of three-component multiple-schedule may allow for a shift toward the study of stimuli that produce reinforcement that matches the stimulation produced by automatically reinforced behavior.

References

Ahearn, W. H., Clark, K. M., DeBar, R., & Florentino, C. (2005). On the role of preference in response competition. *Journal of Applied Behavior Analysis, 38*, 247-250.

Ahearn, W. H., Clark, K. M., Gardenier, N. C., Chung, B. I., & Dube, W. V. (2003). Persistence of stereotypic behavior: Examining the effects of external reinforcers. *Journal of Applied Behavior Analysis, 36*, 439-448.

Ahearn, W. H., Clark, K. M., MacDonald, R. P., & Chung, B. I. (2007). Assessing and treating vocal stereotypy in children with autism. *Journal of Applied Behavior Analysis, 40*, 263-275.

Baer, D. M. (1977). "Perhaps it would be better not to know everything." *Journal of Applied Behavior Analysis, 10*, 167-172.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.

Britton, L. N., Carr, J. E., Landaburu, H. J., & Romick, K. S. (2002). The efficacy of noncontingent reinforcement as treatment for automatically reinforced stereotypy. *Behavioral Interventions, 17*, 93-103.

Cuvo, A. J., May, M. E., & Post, T. M. (2001). Effects of living room, Snoezelen room, outdoor activities on stereotypic behavior and engagement of adult with profound mental retardation. *Research in Developmental Disabilities, 22*, 183-204.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*, 491-498.

Higbee, T. S., Chang, S., & Endicott, K. (2005). Noncontingent access to preferred sensory stimulation as treatment for automatically reinforced stereotypy. *Behavioral Interventions, 20*, 177-184.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (l994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197-209.
Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis, 27*, 131-144.
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.
LeBlanc, L. A., Patel, M. R., & Carr, J. E. (2000). Recent advances in the assessment of aberrant behavior maintained by automatic reinforcement in individuals with developmental disabilities. *Journal of Behavior Therapy and Experimental Psychiatry, 31*, 137-154.
Piazza, C. C., Adelinis, J. D., Hanley, G. P., Goh, H. L., & Delia, M. D. (2000). An evaluation of the effects of matched stimuli on behaviors maintained by automatic reinforcement. *Journal of Applied Behavior Analysis, 33*, 13-27.
Rapp, J. T. (2004). Effects of prior access and environmental enrichment on stereotypy. *Behavioral Interventions, 19*, 287-295.
Rapp, J. T. (2006). Toward an empirical method for identifying matched stimulation for automatically reinforced behavior: A preliminary investigation. *Journal of Applied Behavior Analysis, 39*, 137-140.
Rapp. J. T. (2007). Further evaluation of methods to identify matched stimulation. *Journal of Applied Behavior Analysis, 40*, 73-88.
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.
Rapp, J. T., Miltenberger, R. G., Galensky, T. L., Ellingson, S. A., & Long, E. S. (1999). Functional analysis of hair pulling. *Journal of Applied Behavior Analysis, 32*, 385-388.
Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities, 26*, 527-547.
Rapp, J. T., Vollmer, T. R., St. 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*, 481-501.
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*, 605-620.
Roantree, C. F., & Kennedy, C. H. (2006). Paradoxical effect of precession attention on stereotypy: Antecedent attention as an establishing, not an abolishing, operation. *Journal of Applied Behavior Analysis, 39*, 381-384.
Simmons, J. N., Smith, R. G., & Kliethermes, L. (2003). A multiple-schedule evaluation of immediate and subsequent effects of fixed-time food presentation on automatically maintained mouthing. *Journal of Applied Behavior Analysis, 36*, 541-544.
Vollmer, T. R., Marcus, B. A., & LeBlanc, L. (1994). Treatment of self-injury and hand mouthing following inconclusive functional analyses. *Journal of Applied Behavior Analysis, 27*, 331-344.
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.

Marc J. Lanovaz is currently a doctoral student at McGill University and a researcher-practitioner at the Centre de réadaptation Lisette-Dupras in Montreal, Canada. He received his master's degree in applied behavior analysis from St. Cloud State University in 2007. His current research interests include the assessment and treatment of stereotypy, the functional assessment of problem behavior, and the use of behavior analysis to solve broad social issues.

Sarah E. Fletcher is currently an ABA consultant with the UK Young Autism Project. She received her master's degree in applied behavior analysis from St Cloud State University in 2008. Her current research and clinical interests include treatment of behaviors maintained by automatic reinforcement and discrimination training in children with autism.

John T. Rapp, PhD, is a professor in the behavior analysis and community psychology department at St. Cloud State University. He received his doctoral degree in behavior analysis from the University of Florida in 2003. His current research interests include the assessment and treatment of automatically reinforced behavior, using continuous and discontinuous methods to measure behavior, motivating operations and experimental designs.