Journal of Applied Behavior Analysis

JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2016, **49,** 927–946 NUMBER 4 (WINTER)

EFFECTS OF MULTIPLE VERSUS CHAINED SCHEDULES ON STEREOTYPY AND ITEM ENGAGEMENT

JESSICA D. SLATON

WESTERN NEW ENGLAND UNIVERSITY AND NASHOBA LEARNING GROUP

AND

GREGORY P. HANLEY

WESTERN NEW ENGLAND UNIVERSITY

We evaluated rates of automatically reinforced stereotypy and item engagement for 2 children with autism under multiple and chained schedules in a multielement design. Each schedule included components during which stereotypy was blocked (S–) or allowed (S+), and we used colored cards as schedule-correlated stimuli. We report rates of stereotypy and item engagement during S– and S+ components, as well as the percentage of component time that elapsed before the first instances of stereotypy and item engagement. We observed less stereotypy and more consistent item engagement during chained-schedule sessions, and stimulus control of stereotypy and item engagement was established with the chained schedule. A subsequent concurrent-chains analysis revealed participant preference for the chained schedule. These results highlight the importance of contingent access to stereotypy when therapists attempt to gain stimulus control of stereotypy and increase functional item engagement.

Key words: autism, chained schedules, item engagement, multiple schedules, stereotypy, stimulus control, treatment preference

Stereotypy is described as repetitive and invariant behavior (e.g., hand flapping, body rocking) that does not appear to serve a social function (Rapp & Vollmer, 2005). Chronic, high-rate stereotypy by individuals with autism or other developmental disabilities can be problematic because it may interfere with social activities or skill acquisition, and it may be socially stigmatizing (Cunningham & Schreibman, 2008; Dunlap, Dyer, & Koegel, 1983; Koegel & Covert, 1972; Lanovaz, Robertson, Soerono, & Watkins, 2013; Lovaas, Litrownik, & Mann, 1971). Treatment of stereotypy presents a unique challenge because stereotypy is often maintained by automatic rather than socially mediated reinforcers (Vollmer, 1994), and it is therefore difficult to disrupt the response–reinforcer relation or identify socially mediated reinforcers that are potent enough to compete with the automatic stimulation produced by stereotypy.

An associated challenge in the treatment of stereotypy is that interventions for stereotypy often produce different results across studies and across participants within the same study. For example, Hanley, Iwata, Thompson, and Lindberg (2000) found that simply blocking stereotypy produced increases in appropriate engagement with leisure items for two participants, but was not sufficient to produce these results for a third participant. Potter, Hanley, Augustine, Clay, and Phelps (2013) evaluated similar procedures and found that blocking

This study was conducted in partial fulfillment of the requirements for a PhD in Behavior Analysis from Western New England University by the first author. We thank the staff of Nashoba Learning Group for their support, particularly Kelley Ward, Katie Marshall, Colleen Meagher, Katie Russo, Caitlin MacNeil, Sabrina Congram, Tiffanie Roper, Kim Steele, Mia Morgan, Jill Arruda, and Rachel Rutfield for filming sessions and assisting with data collection.

Address correspondence to Gregory P. Hanley, Western New England University, 1215 Wilbraham road, Springfield, Massachusetts 01119 (e-mail: ghanley@wne.edu).

doi: 10.1002/jaba.345

stereotypy and prompting appropriate engagement alone did not reduce stereotypy. There is evidence that providing access to activities that match the stimulation produced by stereotypy can be effective in reducing stereotypy (Piazza, Adelinis, Hanley, Goh, & Delia, 2000), but there is also evidence that noncontingent access to matched stimulation can produce increases in other forms of untargeted stereotypy (Rapp et al., 2013). Response interruption and redirection (RIRD) is another common stereotypy intervention whose efficacy has been demonstrated by several authors (Ahearn, Clark, McDonald, & Chung, 2007; Ahrens, Lerman, Kodak, Worsdell, & Keegan, 2011; Casella, Sidener, Sidener, & Progar, 2011). However, Carroll and Kodak (2014) and Wunderlich and Vollmer (2015) both demonstrated that different data collection and reporting methods with RIRD (scoring stereotypy outside the intervention time vs. across the entire session) can influence the perceived efficacy of RIRD by overestimating reductions in stereotypy. Thus, in addition to difficulty interrupting the response–reinforcer relation, treating stereotypy can also be challenging because the limited treatment options may not always produce desired results.

When considering the challenges associated with treating automatically reinforced stereotypy, it is also important to consider the possible outcomes of treatment (e.g., reducing stereotypy, eliminating stereotypy, or shifting response allocation). The goal when treating dangerous problem behavior exhibited by persons with autism is often complete elimination of the behavior (e.g., Hanley, Jin, Vanselow, & Hanratty, 2014). However, when treating automatically reinforced stereotypy that is not dangerous (i.e., does not produce tissue damage to self or others), complete elimination may not be a practical, necessary, or even appropriate goal. Consider first that typically developing individuals engage in repetitive and presumably automatically reinforced behaviors (e.g., finger tapping, hair twirling). The occurrence of stereotypy has been reported in typically developing 2-year-olds and preschoolers (Foster, 1998; Leekam et al., 2007; MacDonald et al., 2007) and school-age children (Tröster, 1994), and has been self-reported by college students (Rafaeli-Mor, Foster, & Berkson, 1999). Consider also that complete elimination of these same behaviors for individuals with autism or other developmental disabilities would not be in line with the value of balancing habilitative goals with an individual's right to choose activities (Bannerman, Sheldon, Sherman, & Harchik, 1990) or the right to a therapeutic environment that "imposes the fewest restrictions necessary" (Van Houten et al., 1988, p. 382). Given these two considerations and the limitations of interventions such as RIRD or noncontingent matched stimulation described above, eliminating stereotypy may not be the most appropriate or achievable goal.

Rather than attempting to eradicate stereotypy, a more desirable and practical treatment goal may be to establish stimulus control of stereotypy with stimuli that can easily be presented and removed as the acceptability of engaging in stereotypy changes from context to context (e.g., Brusa & Richman, 2008; Doughty, Anderson, Doughty, Williams, & Saunders, 2007). A second desirable and practical goal would then be to increase engagement in more appropriate social, leisure, or instructional activities so that these behaviors may occur in place of stereotypy during times when stereotypy is not compatible with environmental expectations (Hanley et al., 2000; Lanovaz et al., 2013). Such a shift in response allocation could remove the barriers to learning that stereotypy presents without overly restricting an individual's activities.

Interventions to establish stimulus control of stereotypy and increase appropriate item engagement have been evaluated in two different literatures. Several studies have examined stimulus control of stereotypy via schedule-

correlated stimuli but without measuring or strengthening item engagement (e.g., Cook, Rapp, Gomes, Frazer, & Lindblad, 2014; Conroy, Asmus, Sellers, & Ladwig, 2005; Doughty et al., 2007; Haley, Heick, & Luiselli, 2010; O'Connor, Prieto, Hoffmann, DeQuinzio, & Taylor, 2011; Piazza, Hanley, & Fisher, 1996; Rapp, Patel, Ghezzi, O'Flaherty, & Titterington, 2009). Likewise, some studies have examined contingent access to stereotypy as a reinforcer for item engagement (i.e., completion of academic demands or appropriate engagement with leisure materials) but without the use of correlated stimuli to establish stimulus control (e.g., Charlop, Kurtz, & Casey, 1990; Hanley et al., 2000; Potter et al., 2013). We are not currently aware of any studies that have demonstrated stimulus control of stereotypy as well as strengthened item engagement.

Procedures to establish stimulus control of stereotypy generally involve blocking stereotypy or reprimanding it in the presence of a particular stimulus (the S–) and allowing it in the presence of another stimulus (the S+). Stimulus control of stereotypy affords the benefit of suppressing stereotypy only when necessary and without the need for constant monitoring and redirection that can be difficult for caregivers to provide (Kliebert, Tiger, & Toussaint, 2011). For example, O'Connor et al. (2011) established stimulus control of stereotypy with a young child with autism during a leisure activity (looking at books). Motor stereotypy was blocked while the child looked at books in the presence of a red card; free access to stereotypy was allowed in the presence of a green card. After the duration of the S– component (red card) was systematically increased, stereotypy was also suppressed in the presence of the red card during generalization sessions in the participant's classroom and at a public library.

Many studies that have evaluated stimulus control of stereotypy have done so under a multiple schedule in which, by definition, the S– and S+ components are signaled by correlated stimuli and the changeover between the components is time based. Stereotypy treatment delivered within a multiple schedule has been evaluated with children with autism in the context of motor stereotypy during math instruction (Conroy et al., 2005), vocal stereotypy in a general education classroom (Haley et al., 2010), motor stereotypy while looking at books (O'Connor et al., 2011), and vocal stereotypy while playing with toys (Rapp et al., 2009). Multiple schedules in the context of treatment of stereotypy have also been evaluated with adults with developmental disabilities (Doughty et al., 2007). In each case, the use of a multiple schedule shifted response allocation so that more stereotypy occurred during the S+ component than during the S– component.

It is important to note, however, that a change in response allocation does not necessarily mean that stimulus control has been established. It is possible that stereotypy is instead under direct control of the blocking procedure or reprimand (i.e., stereotypy is under contingency control rather than stimulus control). For example, the individual may engage in stereotypy quickly in the presence of both the S+ and S–, but then refrain from stereotypy during the rest of the S– interval after experiencing the blocking consequence. This would produce lower rates of stereotypy in the presence of the S– relative to the S+, but would also suggest that stereotypy is under contingency control of blocking rather than stimulus control of the S–. If stereotypy is under contingency control, presumably caregivers would need to continue to monitor the individual during the S– to deliver the blocking consequence. An intervention that requires constant monitoring is in direct contrast to one of the main purposes of attempting to gain stimulus control of stereotypy. It is therefore important to know whether changes in stereotypy allocation reflect contingency control or stimulus control.

One way to demonstrate stimulus control is to show that stereotypy occurs to some extent during the S+ and not at all during the S–. This outcome has not yet been reported in studies that have evaluated multiple schedules. Doughty et al. (2007) addressed the question of contingency control versus stimulus control by measuring latency to the first instance of stereotypy in each component of the multiple schedule. These authors found that much higher percentages of component time elapsed before the first instance of stereotypy during the S– component than during the S+ component (70% or more of S– component and 30% or less of S+ component elapsed), which they interpreted as indicative of stimulus control. Rapp et al. (2009) applied the same procedures described by Doughty et al. (2007) and concluded that stimulus control had been established for one of three participants. Although greater latencies to engage in stereotypy during the S– are indeed indicative of stimulus control, these outcomes still necessitate constant monitoring during the S– to implement the correlated procedure.

Increasing appropriate engagement in social, leisure, or instructional activities is also an important goal of interventions that target stereotypy. However, measures of appropriate engagement are typically not reported in studies that examine stimulus control of stereotypy. For example, Cook, Rapp, Gomes, Frazer, and Lindblad (2014) noted that a lack of appropriate alternative activities during the S– component was a limitation of their procedure; these authors recommended that additional research evaluate stimulus control of stereotypy concurrently with measures of appropriate engagement. Some authors have reported providing specific appropriate activities during S– components (e.g., O'Connor et al., 2011; Rapp et al., 2009), but they did not collect data on engagement with those activities. As an exception, Conroy et al. (2005) reported data on rates of engagement during math instruction, but these data showed that engagement did not change relative to baseline. Thus, it is unclear whether multiple schedules that target automatically reinforced stereotypy strengthen repertoires of appropriate behavior such as functional item engagement.

Appropriate engagement has been strengthened in individuals with chronic stereotypy when contingent access to stereotypy has been arranged as a consequence for engagement. This procedure generally involves presenting relevant academic or leisure tasks, prompting appropriate responding while restricting access to stereotypy, and then allowing stereotypy contingent on some initially minimal and then gradually increased amount or duration of appropriate responding. For example, Charlop et al. (1990) described an intervention in which children with autism were given the opportunity to engage in stereotypy after correct responses during academic instruction. The percentage of correct responses was higher for each child during sessions in which access to stereotypy was provided after correct responses than in sessions in which other (edible) reinforcers were provided. Contingent access to stereotypy has also been shown to increase appropriate leisure-item engagement in teenagers and adults with autism (Hanley et al., 2000; Potter et al., 2013). In addition, Potter et al. (2013) showed that the contingency between item engagement and stereotypy access was critical to the desirable outcomes by comparing conditions in which contingent access to stereotypy was and was not provided. Although these studies resulted in large decreases in stereotypy and increases in appropriate responding, stereotypy was not eliminated for any participant during the period in which engagement was expected. The absence of overt schedule-correlated stimuli from the procedures may have contributed to the low but persistent rates of stereotypy.

Studies on stimulus control of stereotypy and studies on contingent access to stereotypy

each provide separate independent variables for producing two desirable outcomes relevant to treating stereotypy: gaining stimulus control of stereotypy and strengthening appropriate behavior, respectively. An intervention that combines schedule-correlated stimuli and contingent access to stereotypy may potentially produce both desired outcomes. This sort of intervention may be best characterized as a chained schedule in that there are two components, each with a correlated stimulus, and alternation between components is dependent on performance in at least one of the components (Bejarano & Hackenberg, 2007; Kelleher & Gollub, 1962). In other words, access to the S+ component is noncontingent in a multiple schedule but is contingent in a chained schedule. This means that in a multiple schedule, stereotypy is suppressed during the S– component only to the extent that blocking either interrupts the response–reinforcer relation or functions as a punisher (e.g., Ahrens et al., 2011); there are no other programmed contingencies to weaken stereotypy or strengthen any other response. It is possible that a chained schedule may therefore be more effective than a multiple schedule in suppressing stereotypy and strengthening item engagement because a chained schedule includes response blocking during the S– component in addition to programmed contingencies that operate on both stereotypy and item engagement. We conducted an evaluation of this synthesized intervention in the current study. We first used a multielement design to compare rates of stereotypy and appropriate item engagement under multiple and chained schedules with two children with autism. Both schedules included response blocking of stereotypy during the S– component. We then arranged a concurrent-chains analysis, similar to Potter et al. (2013), to identify the children's preference for stereotypy treatments arranged within multiple or chained schedules.

METHOD

Participants and Settings

We recruited participants from a private day school for children with autism, and all sessions took place in classrooms at the school. Four criteria were used to accept participants into the study: (a) reducing stereotypy and increasing functional item engagement were objectives in the participant's individualized education plan; (b) problem behaviors such as aggression, self-injury, property destruction, or some combination, did not occur on a regular basis; (c) the participant already had some independent play or work skills in his or her repertoire; and (d) results of a functional analysis revealed that the participant's stereotypy was not sensitive to socially mediated consequences. Two participants met these criteria; their data are reported here.

Mark was an 11-year-old boy who had been diagnosed with autism; he communicated vocally in two- to three-word phrases or short sentences that consisted primarily of mands for preferred items and activities. Molly was an 18-year-old young woman who had been diagnosed with autism and a seizure disorder; she communicated with a speech-generating device (SGD) to emit one- and two-word mands for preferred items and activities. Both participants engaged in high levels of motor stereotypy (topographies described below). Mark's and Molly's teachers reported that their stereotypy interfered with all areas of instruction (e.g., independent leisure activities, vocational tasks, one-on-one instructional sessions, social interaction, daily living activities, transitioning between environments). In particular, Mark's teachers expressed concern that he engaged in stereotypy when directed to select independent leisure activities and therefore could not play appropriately by himself during free time. Molly's teachers were most concerned about her stereotypy during vocational tasks because it impeded her completion of the work and

therefore reduced her chances of accessing employment opportunities as an adult. Some topographies of Molly's stereotypy were reportedly dangerous depending on the context in which they occurred (e.g., leaping and twirling in a crowded classroom). Several minor accidents related to her stereotypy had been reported (e.g., tripping or striking another person while jumping).

Sessions for Mark were conducted in a small play area in his classroom; sessions for Molly were conducted at a table in her vocational classroom. Participant assent was gained before each session by asking each if they were ready to begin working. An affirmative response included any one of the following: head nodding, saying "yes," saying "okay," or coming to sit down in the designated area. A rejection response included any one of the following or a combination thereof: saying "no thanks," shaking the head implying "no," or using an SGD to reject participation (Molly). Mark never declined any sessions; on rare occasions Molly used her SGD to say "stop" or "all done" when asked to participate in sessions; these requests were honored.

Response Definitions, Measurement, and Interobserver Agreement

Each session was videotaped and then scored using a laptop computer equipped with data-collection and analysis programs. We scored the following measures for both participants in each session: stereotypy count and item-engagement count (both reported as responses per minute) during S– and S+ components, and the onset and offset of each component. Motor stereotypy was scored by counting each movement (e.g., each clap, each jump) for movements that could be individually counted, or by counting each set of movements for actions that occurred too rapidly to be counted individually (e.g., each set of head shakes or body bounces separated by 1 s). We scored item engagement by counting each discrete item manipulated correctly (e.g., each puzzle piece placed in the puzzle or each newsletter folded) in real time as the participant completed the response. We determined responses per minute of stereotypy and item engagement by dividing the total count by cumulative duration in minutes of each component. We determined latency measures by calculating percentage of component time elapsed before the first instance of stereotypy or item engagement (e.g., Doughty et al., 2007). Specifically, we divided the latency in seconds to the first instance by the cumulative duration (in seconds) of each component. If stereotypy or item engagement did not occur at all during the S– or S+ components that session, we scored the percentage of component time elapsed as 100%. Timers and iPhones were used to time the duration of each S– and S+ component.

Topographies of stereotypy were defined individually for Mark and Molly based on prior observations of each participant by the first author and discussions with each participant's lead teacher. Motor stereotypy for both participants consisted of repetitive movements of the arms, legs, head, or whole body as well as repetitive movements with objects, such as tapping or shaking toys and work materials. Repetitive arm and leg movements included actions such as knocking, tapping, clapping, waving hands in front of face, wringing hands, raising and lowering arms, shaking arms, or kicking. Repetitive head movements included actions such as shaking head from side to side (as if emphatically saying "no") and tilting head back and forth while squinting. Whole-body movements included rocking, bouncing, and jumping. During all sessions for Mark, several toys were present that were regularly available during his day and with which he reportedly engaged in high levels of stereotypy (e.g., knocking on the toys or tapping them against his body). Stereotypy was scored each time Mark shook,

knocked on, or tapped a toy or tapped it against his body. For Molly, vocational work materials were present during all baseline and treatment sessions. Stereotypy with work materials consisted of picking up a handful of items and sifting them through her fingers, or repeating an action with an item. For example, when she packaged items into a bag, Molly often picked up the bag, put it down, picked it up again, and repeated this several times before placing items in the bag. Each repeated action was scored as an instance of stereotypy.

Independent leisure or vocational work activities were also selected individually based on teacher reports of activities that Mark and Molly had reliably completed without prompts, but during which stereotypy also occurred. Of these, only activities that involved some product completion were included (e.g., doing puzzles, stringing beads, folding newsletters). This criterion was applied to ensure that item engagement for each task could be reliably counted. Mark's appropriate leisure activities were puzzles, beads to string, a set of wooden blocks with photographs showing specific block structures to build, and binders with pages of removable Velcro words and pictures to match together. Item engagement was defined for Mark as follows: each puzzle piece placed correctly in the frame; each bead placed on the string; each block placed correctly according to the current block picture; each word placed on the correct picture. Molly's activities were office-related vocational tasks: folding newsletters in half, counting and packaging small items into a plastic zip bag, removing staples from documents, and stuffing envelopes with brochures. For counting and packaging items, Molly was given the written numeral 1 or 2, a set of small round items (metal washers of different sizes), and a set of plastic zip bags. She placed the correct number of items on the numeral card and then placed the items in a bag. Each newsletter folded, staple removed, and brochure placed in an envelope was counted individually. For counting and packaging items, each item placed on the numeral card was counted. Additional items placed on the numeral card in excess of the designated quantity were not scored. When Molly then picked up the items and placed them in the bag, another instance of item engagement was scored.

A second observer viewed session videos and collected data for the purposes of calculating interobserver agreement. Data were collected using a laptop computer equipped with a data-collection program to score instances of stereotypy, item engagement, and the durations of the S+ and S– components. Agreement was calculated by dividing session data into 10-s intervals and dividing the number of agreements per interval by the number of agreements and disagreements per interval and converting the result to a percentage (i.e., partial agreement within intervals). Agreement was scored for at least 20% of sessions in each condition for each participant. Agreement for stereotypy averaged 88% for Mark (range, 82% to 94%) and 88% for Molly (range, 81% to 97%), for item engagement averaged 95% for Mark (range, 80% to 100%) and 96% for Molly (range, 90% to 100%), and for component duration averaged 95% for Mark (range, 85% to 99%) and 93% for Molly (range, 86% to 96%).

Design

We conducted sessions 3 to 5 days per week in blocks of two to six sessions. A multielement design was used during the functional analysis and treatment-efficacy analysis. Session order was determined randomly in advance each day, with the requirement that no condition be presented for more than two consecutive sessions and that each condition be presented an equal number of times. During the treatment analysis, the duration of each S– component in the multiple schedule was yoked to the duration of each S– component in the previous chained

schedule session. A concurrent-chains design with forced-choice and free-choice trials was used to identify participant preference after the treatment analysis.

Functional Analysis

We completed a functional analysis of stereotypy for both Mark and Molly using procedures similar to those described by Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994). The attention condition included a brief reprimand as well as response blocking contingent on each instance of stereotypy (e.g., "hands down, please" while prompting the participant to put his or her hands back down). Beyond revealing persistence of stereotypy in the absence of social consequences, analysis of social-positive and social-negative contingencies was important because both the multiple and the chained schedules involved blocking stereotypy, which provided physical attention as well as brief escape from a task contingent on stereotypy.

Comparison of Multiple and Chained Schedules

Baseline. Mark's teachers reported that even though he had mastered independent leisure activities such as doing puzzles or building with blocks, when presented with these materials during free time he would shake and tap these and other items. Thus, we designed Mark's baseline sessions to include a variety of mastered leisure tasks as well as the presence of items with which he reportedly engaged in high rates of stereotypy, and we conducted these sessions in the play area where he typically spent his break time. The play area consisted of a bookshelf with multiple leisure activities and a space for him to sit on the carpet in front of the shelf. Molly's teachers reported that even though she had mastered several office-related vocational tasks, she still required a high level of supervision during these tasks because she engaged in stereotypy with the task materials or after she completed a work product. We therefore designed Molly's baseline sessions to include a variety of office-related vocational tasks, and we conducted her sessions in the vocational office training room.

All sessions lasted 6 min. Each baseline session began with the participant seated in front of his or her materials and with a colored card placed to the left. A green card and a black card were used as one set of stimuli, and a white card and a yellow card were used as the other set of stimuli for both participants. Cards were not assigned to the multiple or chained schedule until after baseline was complete. The cards were laminated pieces of construction paper (10 cm by 15 cm) taped together to create a thick card with a different color on each side. Ten tokens (round plastic discs) were attached with Velcro to one card to evaluate whether differentiated responding occurred in the presence of the tokens before their contingent delivery in the chained schedule.

The experimenter instructed the participant either to play with the leisure activities (Mark) or to do some work (Molly). The experimenter flipped the colored card over to the other side at the end of every minute and provided brief praise for whatever action the participant happened to be engaged in at that moment. If the participant was engaged with the task materials, praise was provided. If the participant was not engaged, he or she was praised for remaining in the designated area, sitting calmly, or remaining upright (Mark attempted to lie down and nap during baseline). There were no programmed consequences for stereotypy.

Treatment. Stimulus sets were randomly assigned to either the chained schedule or the multiple schedule. Within each stimulus set, one color was assigned to the component during which stereotypy was blocked (S–) and the other to the component during which stereotypy was allowed (S+). The color associated with the highest rates of stereotypy or an

increasing trend (or both) during baseline was assigned to the S– component for each schedule. Each session began with the experimenter holding up the card for the S– component, prompting the participant to touch the card, and giving a verbal instruction. For Mark, the instruction was "We're going on [color] now, so play with the toys." For Molly, the instruction was "We're going on [color] now, so do the work." Blocking during the S– component for each schedule was implemented by having the experimenter physically interrupt each instance of stereotypy by placing her hands on the participant's hands, arms, shoulders, or head (depending on the topography of stereotypy) for 1 to 2 s. If the instance of stereotypy was completed before the experimenter blocked it (this rarely occurred), 1 to 2 s of physical contact with the participant was still provided to prevent the occurrence of an additional instance of stereotypy. Thus, both attempted and completed instances of stereotypy were blocked and were scored during the S– interval.

Chained schedule. During the S– component of the chained schedule, participants were required to engage in a certain number of leisure or vocational responses in succession without stereotypy before changeover to the S+ component occurred. This response requirement was illustrated with tokens attached to the card correlated with the S–. Item engagement was not prompted. If Mark or Molly made an error with a task (e.g., placing a puzzle piece in the wrong spot or folding a newsletter incorrectly), the experimenter corrected the error and did not deliver a token for that response. One token was initially delivered after each instance of engagement, and this schedule was thinned every several sessions until reaching the terminal values of fixed-ratio (FR) 10 FR 4 (a total of 40 responses) for Mark and FR 10 FR 2 (a total of 20 responses) for Molly. For example, at first, one token was provided for each puzzle piece that Mark placed in the puzzle frame, and nine tokens were required before changeover to the S+ occurred (FR 9 FR 1, a total of nine responses). Next, two responses were required to earn one token, and 10 tokens were required before changeover to the S+ occurred (FR 10 FR 2, a total of 20 responses). Eventually, one token was delivered for every four responses (i.e., every four puzzle pieces), and 10 tokens were required. Changeover to the S+ component therefore always occurred after a chain of successive appropriate responses, such as connecting 40 puzzle pieces or folding 20 newsletters without stereotypy. The duration of the S+ component was 30 s across all sessions for both participants. However, the duration of the S– component in the chained schedule was dependent on the amount of time it took the participant to complete his or her required responses without stereotypy. As this response requirement increased, the average duration of the S– component also increased as an artifact of the change in response requirement.

When an instance of stereotypy occurred, the response requirement was reset by removing all tokens after blocking was implemented. If the participant had not yet earned any tokens, only blocking was implemented. Component changeover occurred when all tokens had been earned. The experimenter turned over the card to reveal the color correlated with the S+, prompted the participant to point to it, and said, "We're going on [color], so you can do your own thing now." The experimenter then moved approximately 0.6 m away from the participant. During the first three sessions for each participant, the experimenter also provided instructions and models indicating what the participant was now allowed to do. For example, for Mark, the experimenter said, "We're going on black now, so you can do your own thing. You can bounce and shake like this if you want [model bouncing] and you can shake the toy." For Molly, the experimenter said, "We're going on green now, so you can do your own thing. You can get up and twirl

like this if you want [model twirling] and you can jump and shake your head." There were no programmed consequences for stereotypy or item engagement during the S+ component. Each S+ component lasted 30 s, after which the experimenter approached the participant and implemented the next S– component by repeating the changeover procedure (hold up the card, prompt the participant to touch it, provide the verbal instruction). Each chained schedule session lasted until three S+ components had been experienced. However, if three S+ components had not been experienced after 30 min, the session was terminated after the next S+ component past the 30-min mark. On one occasion, Molly's session was terminated after only two S+ components because the 30-min cap was reached. Average session duration was 15 min for Mark and 13 min for Molly.

Multiple schedule. Multiple-schedule sessions were identical to the chained-schedule sessions except that component changeover occurred after a fixed amount of time had elapsed regardless of any item engagement or instances of stereotypy, and no tokens were present. Durations of each S– component were yoked to the previous chained-schedule session and were presented in the same order in which they had occurred during the previous chained session to ensure that Mark and Molly both experienced an equal amount of time in the S– components of each schedule. If an instance of stereotypy occurred during the S– component of the multiple schedule, blocking was implemented, and no other stimulus changes occurred. The duration of each S+ component remained 30 s.

Preference analysis. After the efficacy analysis for multiple versus chained schedules had been completed, a concurrent-chains analysis was conducted to identify preference for treatments (Hanley, 2010). Initial-link stimuli were developed for the multiple schedule, the chained schedule, and a control condition in which stereotypy was blocked but no leisure or work materials were present. The control condition was included to distinguish indiscriminate responding from lack of preference. The initial-link stimuli for each schedule were pictures of the schedule-correlated cards next to each other (e.g., a picture of the white and yellow cards next to each other for the multiple schedule; a picture of the green and black cards with tokens next to each other for the chained schedule). The initial-link stimulus for the control condition was a picture of an empty bookshelf to indicate that no materials were available.

Initial-link stimuli were presented on a strip of poster board, evenly placed and secured with Velcro. Before each trial, the participant was asked to look at the choices and then was instructed to "pick the one you want to do." A response was scored when the participant removed an initial-link stimulus from the poster board and placed it in the experimenter's outstretched hand. The experimenter then said the participant's choice and began the terminal-link session (e.g., "Okay, you chose white and yellow; let's go do that."). A series of 12 free-choice trials were presented to Mark over the course of 2 days. Molly was presented with five free-choice trials and then a series of forced-choice trials in which all terminal-link stimuli were presented and she was instructed to "pick this one" (with a gesture to a specific stimulus) to ensure that she experienced each terminal-link session. These forced-choice trials were conducted approximately 30 min apart over the course of 2 days until each terminal link had been experienced twice. Free-choice trials were then presented to Molly again. The same criteria for session duration used in the efficacy analysis were also applied to terminal-link sessions: A session ended when three S+ components were experienced or when 30 min had elapsed, and multiple-schedule sessions were yoked to the most recent chained-schedule session. No terminal-link sessions reached the 30-min session cap.

RESULTS

Functional Analysis

Both Mark and Molly engaged in the highest rates of stereotypy during the alone condition relative to any other condition, suggesting that their stereotypy was not sensitive to socially mediated consequences (see Figure 1).

Comparison of Multiple and Chained Schedules

Rates of stereotypy. Stereotypy occurred at high and variable rates during baseline for both participants, and there was no change in rates given the different-colored cards. Mean stereotypy during baseline was 22 responses per minute for Mark and 5 responses per minute for Molly. Stereotypy was reduced in the S– components of both schedules for both participants during the treatment comparison, but we observed a greater reduction during the S– component of the chained schedule than in the multiple schedule (see Figure 2, top). Mean treatment rates of stereotypy during the S– component for Mark were 0.2 responses per minute in the chained schedule and 2 responses per minute in the multiple schedule (reduced from means of 24 and 23 responses per minute in baseline, respectively). Mean treatment rates of stereotypy during the S– component for Molly were 0.8 responses per minute in the chained schedule and 2.7 responses per minute in the multiple schedule (reduced from means of 6.4 and 3.7 responses per minute in baseline, respectively). These data represent reductions in stereotypy of 99% in the S– component of the chained schedule and 91% in the S– component of the multiple schedule for Mark, and reductions in stereotypy of 88% in the S– component of the chained schedule and 46% in the S– component of the multiple schedule for Molly. Stereotypy remained at high and variable rates during the S+ component of both schedules for both participants (see Figure 2, second panels).

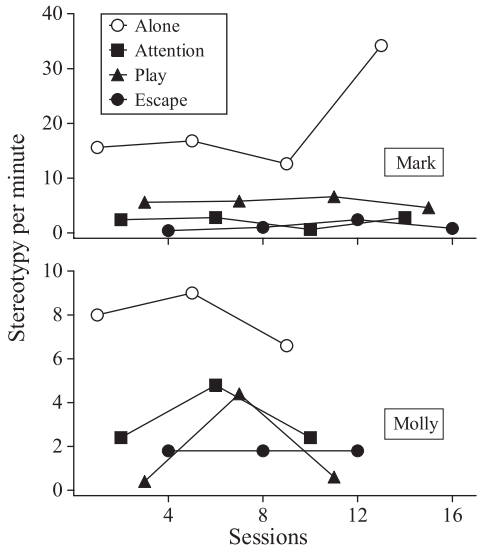

Figure 1. Functional analyses of stereotypy for Mark and Molly, indicating that stereotypy was not reinforced by socially mediated consequences.

Rates of item engagement. Mark did not demonstrate any item engagement during baseline; Molly demonstrated a mean of 1 response per minute of item engagement during baseline, but she also engaged in stereotypy while performing her work tasks.

Item engagement increased during the S– component of both schedules for both participants during the treatment comparison, but we observed more consistent engagement and higher rates of engagement during the S– component of the chained schedule than in the multiple schedule (see Figure 2, third panels). Mean rates of item engagement during the S– component for Mark were 5.4 responses per minute in the chained schedule and 3.2 responses per minute in the multiple schedule (both increased from 0 in baseline). Mean rates of item engagement during the S– component for Molly were 4.4 responses per minute in the chained schedule and 3.3 responses per minute in the multiple schedule (increased from means of 0 and 1.3 responses per minute in baseline, respectively). Item engagement was observed during 100% of chained-schedule sessions for

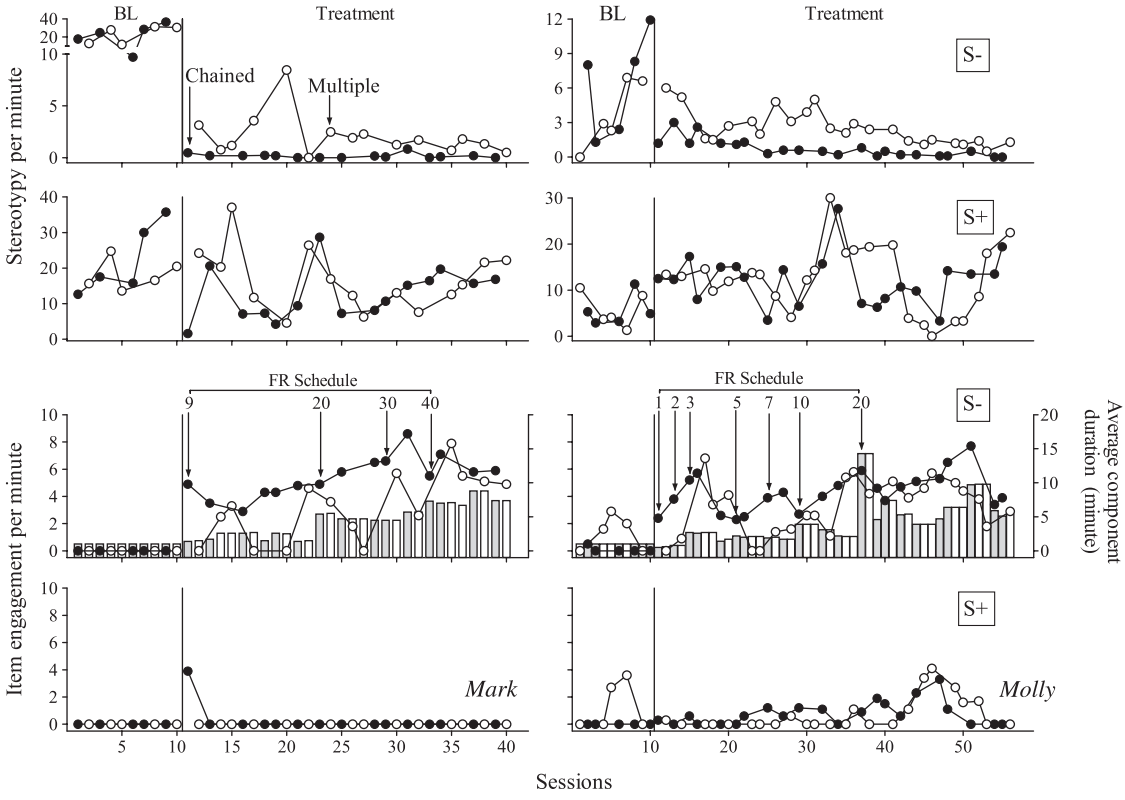

Figure 2. Rates of stereotypy and item engagement for Mark (left) and Molly (right) during the multiple versus chained schedule comparison. Gray and white bars correspond to the secondary *y* axis, indicating average component duration for the S– component; S+ component duration was always 30 s and is not depicted on the graph.

both participants, compared to only 73% of multiple-schedule sessions for Mark and 88% of multiple-schedule sessions for Molly.

Item engagement during the S+ component was low in both schedules for both participants (see Figure 2, bottom). Mark engaged in some instances of item engagement during the S+ component of the first chained session, but then exclusively engaged in stereotypy during the S+ components of all remaining sessions. For Molly, beginning with Session 22 in the chained schedule and Session 43 in the multiple schedule, we observed some item engagement during S+ components. That is, during some sessions, when component changeover occurred and Molly was told, "You can do your own thing now," she continued to engage with the work materials for a portion of the S+ component while she also engaged in stereotypy.

By the end of the comparison, Mark was completing 40 responses per S– component in the chained schedule (for a total of 120 responses per session), and his average S– component duration had increased from 2.1 to 7.6 min. Molly was completing 20 responses per S– component in the chained schedule (for a total of 60 responses per session) by the end of the comparison, and her average S– component duration had increased from 0.6 to 7 min.

Latency to stereotypy. During baseline, stereotypy tended to occur early in the interval before a significant portion of component time had elapsed (*M*s = 4% for Mark and 9% for Molly), indicating that none of the colored

cards exerted inhibitory stimulus control of stereotypy. During the treatment comparison, we observed higher latencies to the first instance of stereotypy in the S– component of the chained schedule (see Figure 3, top). For Mark, we observed complete suppression of stereotypy during the S– component of the chained schedule during 33% of sessions, compared to only 7% of sessions (one session) in the multiple schedule. For Molly, latency to the first instance of stereotypy during the S– component of the chained schedule slowly increased across sessions until stereotypy was completely suppressed during the S– component in the final two sessions of the chained schedule. By contrast, latency to the first instance of stereotypy during the S– component of the multiple schedule remained low for both participants (*M*s = 20% of S– component time elapsed for Mark and 2.8% for Molly). These data suggest that inhibitory stimulus control of stereotypy was more strongly established for both Mark and Molly with the S– in the chained schedule than with the S– in the multiple schedule.

Latency to item engagement. We observed shorter latencies to the first instance of item engagement in the S– component of the chained schedule than in the multiple schedule (see Figure 3, third panels). For Mark, the mean component time elapsed before the first instance of item engagement in the S– component was 8% for the chained schedule and 31% for the multiple schedule. For Molly, the

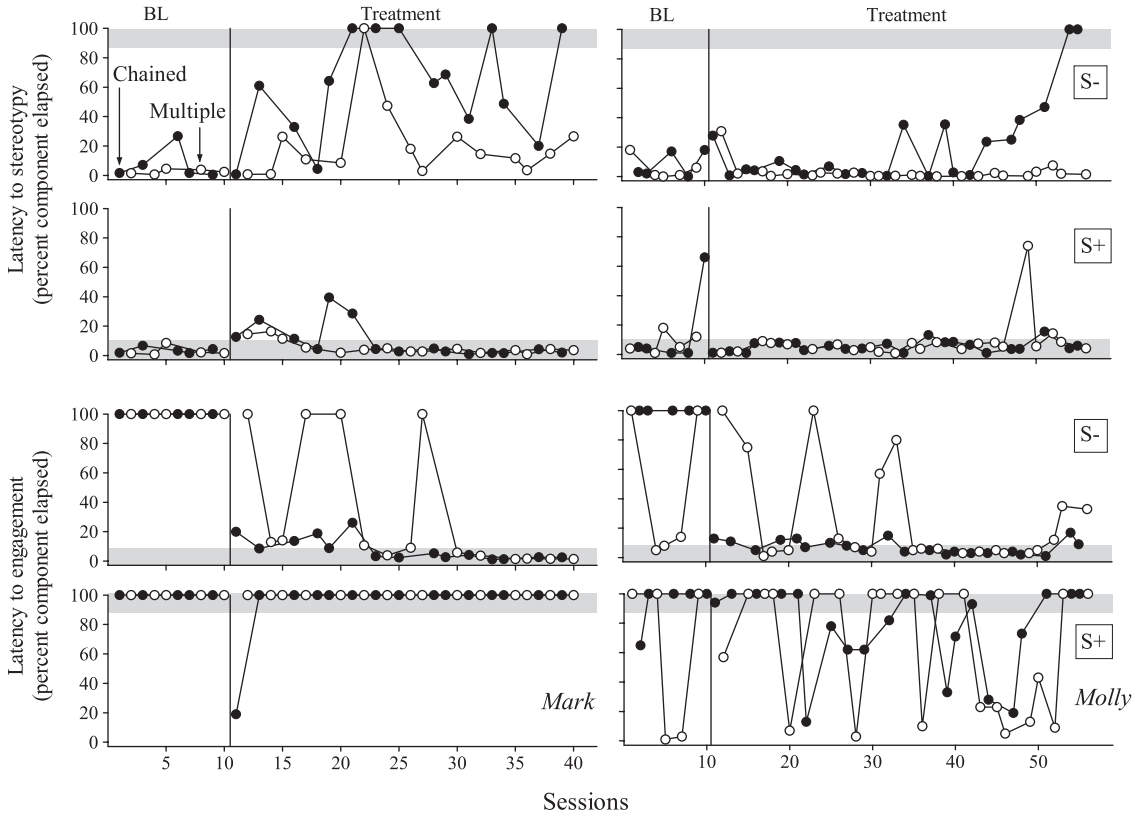

Figure 3. Latency to the first instance of stereotypy and item engagement for Mark (left) and Molly (right) during S– and S+ components of the multiple versus chained schedule comparison.

mean component time elapsed before the first instance of item engagement during the S– component was 7% for the chained schedule and 25% for the multiple schedule. In other words, during the S– component of the chained schedule, Mark and Molly both quickly engaged their leisure or work materials but tended to hesitate or not engage the materials during the S– component of the multiple schedule. This response pattern suggests that the S– component in the chained schedule gained stronger stimulus control of item engagement.

Within-session patterns. Figure 4 highlights some within-session response patterns we observed, using example sessions from Molly. These sessions were chosen because they illustrate a particularly important difference in response patterns that can occur between multiple and chained schedules. In the chained schedule, stereotypy was often allocated to the S+ component only, and did not occur very often during the S– component. By contrast, in the multiple schedule, although stereotypy occurred at a higher rate during the S+ component, it also occurred throughout the S– component. For this particular pair of sessions, we did not observe significant differences in overall amount of stereotypy (38 instances during the chained session and 51 instances during the multiple session), but there was a noticeable difference in how stereotypy was allocated to the S– and S+ components. Finally, Molly consistently engaged the items during the S– component of the chained schedule, but during the multiple schedule, she did not engage the items until the final S– component.

Preference results. The concurrent-chains analysis indicated that both participants selected the chained schedule over the multiple

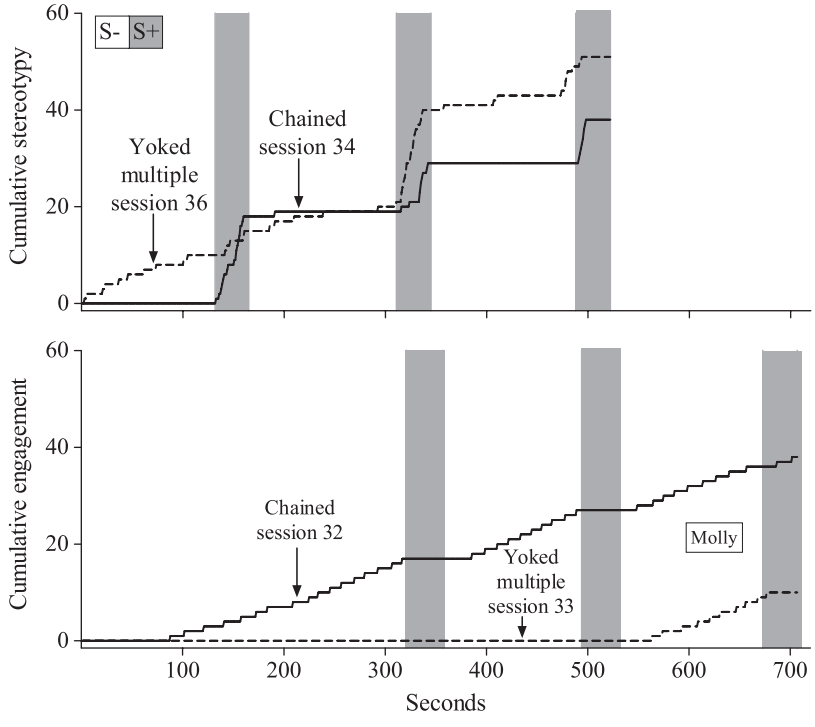

Figure 4. Cumulative records depicting some within-session response patterns for stereotypy and item engagement for Molly.

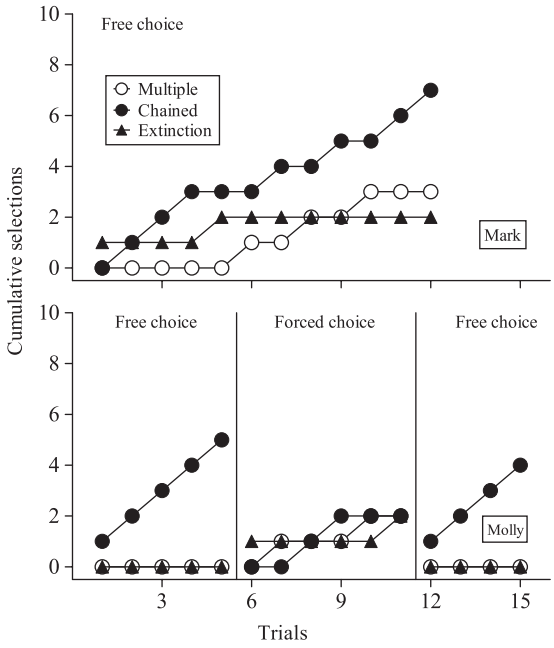

Figure 5. Results of concurrent-chains preference analysis.

schedule and a control condition in which stereotypy was blocked and no leisure or work materials were available (see Figure 5). Mark selected the chained schedule seven times, the multiple schedule three times, and the control condition twice. For Molly, after selection of the chained schedule for five consecutive trials, a series of forced-choice trials were presented to ensure that she experienced each terminal-link session before free-choice trials were presented again. Molly continued to allocate her responding exclusively to the initial-link stimulus associated with the chained schedule as she had during the initial round of free-choice trials.

DISCUSSION

Better stimulus control of stereotypy and functional item engagement was produced with chained schedules than with multiple schedules. These two outcomes have not previously been demonstrated together. These results extend the work of Cook et al. (2014), Doughty et al. (2007), O'Connor et al. (2011), and Rapp et al. (2009) by measuring and increasing appropriate item engagement while gaining stimulus control of stereotypy. We believe these data also emphasize the importance of contingent access to stereotypy when one attempts to establish stimulus control of automatically reinforced stereotypy.

Both Mark and Molly engaged in lower rates of stereotypy and higher rates of item engagement during the S– component of the chained schedule than in the S– component of the multiple schedule. Both schedules shifted rates of stereotypy and item engagement, but latencies to stereotypy were consistently higher during the S– component of the chained schedule for both participants. The complete suppression of stereotypy during the S– component was observed more often in the chained schedule than in the multiple schedule, suggesting that stereotypy in the chained schedule came under contingency control and then inhibitory control of the S–, whereas stereotypy in the multiple schedule remained somewhat under contingency control of blocking. By the end of the treatment comparison, both participants consistently engaged with their leisure or vocational tasks for total session durations of 20 to 25 min in the chained schedule with little to no stereotypy and with very little contingent access to stereotypy provided throughout the session (a total of 90 s of contingent access, divided among three S+ intervals of 30 s each). The primary implication of these findings is that signaling an appropriate time and place to engage in stereotypy will likely lower the rate of stereotypy in conditions in which stereotypy may be undesirable; however, this strategy is not likely to lead to practical stimulus control of stereotypy unless the chance to engage in stereotypy is earned for refraining from stereotypy and for engaging in another contextually relevant behavior. There is some evidence to suggest that simply blocking stereotypy will

produce an increase in item engagement (Hanley et al., 2000); we did not observe this to be the case in the multiple schedule when stereotypy was blocked and no contingency was placed on item engagement. Instead, our findings were similar to those reported by Potter et al. (2013), in which differential reinforcement was necessary to produce changes in item engagement. Our findings are also consistent with the conclusions drawn by Lanovaz et al. (2013) in their review of 60 studies on stereotypy reduction. These authors concluded that strengthening specific appropriate responses is a critical component of stereotypy intervention because effective reduction of stereotypy is not likely to be observed otherwise.

There are several possible reasons why the chained schedule was more effective than the multiple schedule in reducing stereotypy, increasing item engagement, and establishing stimulus control of both these responses. First, the chained schedule included a contingency in which access to a known reinforcer (stereotypy) was provided for appropriate responses and the omission of stereotypy. This directly reinforced a chain of appropriate item engagement that did not include stereotypy. By contrast, time-based alternation of schedule-correlated stimuli in the multiple schedule may strengthen whatever response is occurring at the moment the changeover occurs. Thus, in the multiple schedule, access to stereotypy during the S+ component sometimes followed periods in which no item engagement had occurred during the S–component. Item engagement was therefore not directly and consistently reinforced in the multiple schedule. Second, access to stereotypy during the S+ component of the multiple schedule sometimes followed periods in which stereotypy had occurred during the S– component, which meant that stereotypy during the S– component may have been inadvertently reinforced. This inadvertent reinforcement of stereotypy did not occur during the chained schedule.

It is also possible that the chained schedule was more effective than the multiple schedule because it included a response cost (i.e., all tokens were removed after stereotypy in the S– condition). Thus, the chained schedule included contingent access to stereotypy as well as a response cost for engaging in stereotypy. These procedures may be individually responsible for the effects observed or they may have interacted to establish stimulus control of stereotypy and increase item engagement. Regardless of whether these results were the product of individual or interactive effects, Mark's and Molly's efficacy data should be interpreted through the lens of the specific consequences (or lack thereof) that were available during each compound schedule. The key difference between the multiple and chained schedules was the presence of contingent access to stereotypy and a response cost in the chained schedule and the absence of any such contingencies in the multiple schedule. One potential area for future research may be to evaluate the role of response cost in the chained schedule by comparing chained schedules with and without response cost; a second may be to determine whether chained schedules with contingent access to a different reinforcer (e.g., a preferred item or activity) are equally efficacious as chained schedules with contingent access to stereotypy.

Contingent access to stereotypy may also be responsible for the chained schedule being preferred over the multiple schedule. Children's preference for contingent rather than noncontingent reinforcement has been frequently demonstrated with socially mediated reinforcers (e.g., Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997; Luczynski & Hanley, 2009, 2014), even when more reinforcement is available noncontingently (Luczynski & Hanley, 2010). Our results are also consistent with those of Potter et al. (2013), who found that three adolescents with autism preferred contingent access to stereotypy over conditions in

which all stereotypy was blocked, and that two of the three preferred the contingent access condition to one in which they could have engaged in stereotypy uninterrupted (the third participant showed indifference). The efficacy and preference data of Potter et al., along with our own, provide clear support for a balanced approach to addressing stereotypy rather than the more extreme treatments that either consistently restrict or punish stereotypy or allow the individual to engage in stereotypy uninterrupted. These very different approaches have both been shown to be ineffective and nonpreferred by the children who experience these treatments. Our findings and those of Potter et al. also illustrate that arranging contingencies between appropriate behavior and relevant reinforcement is not only effective for addressing socially mediated problem behavior (e.g., Hanley et al., 1997) and preferred by those who experience the contingencies but also is effective and preferred by those who engage in automatically reinforced and noninjurious problem behavior.

A second potential reason that both participants preferred the chained to the multiple schedule may be the additional conditioned reinforcement during the chained schedule. For instance, when reinforcement was delivered contingent on a specific chain of behavior (e.g., folding 20 newsletters or building a tower with 20 blocks), each response in the chain may have served as a conditioned reinforcer (Kelleher & Gollub, 1962). The tokens provided after each response also may have served as conditioned reinforcers (Hackenberg, 2009).

An interesting extension of these results may be to evaluate stereotypy and item engagement within a chained schedule with participants who display different patterns of stereotypy during their pretreatment functional analysis. Hagopian, Rooker, and Zarcone (2015) outlined three different subtypes of automatically reinforced self-injury based on response patterns during standard functional analyses and found that individuals within each subtype tended to respond differently to treatment. Hagopian et al. noted that high rates of automatically reinforced self-injury in the alone condition and relatively low rates in the play condition (Subtype 1) indicate that the behavior is affected by other reinforcers (e.g., leisure items, social interaction) available in the environment, and treatment based on providing these alternative reinforcers is therefore more likely to be effective. Mark's and Molly's functional analysis response patterns are consistent with Subtype 1, and their treatment data are consistent with the predictions made by Hagopian et al. However, the participants in Hanley et al. (2000) all displayed response patterns consistent with Subtype 2, in which responding was elevated across all conditions, but similar treatment components were effective in reducing their stereotypy and increasing item engagement (e.g., blocking, prompting, and differential reinforcement). It may be useful to evaluate treatment of stereotypy delivered within a chained schedule across participants whose functional analyses show different response patterns, because there is some indication that differential reinforcement-based treatments for stereotypy can be effective regardless of response pattern during the functional analysis.

One limitation of the current study is that participants demonstrated some functional item engagement with the leisure or vocational tasks before intervention, although this was reported to occur at low rates. It is unclear how the efficacy of the chained schedule for increasing item engagement may be affected for participants who do not currently have these skills in their repertoire. Prompts for item engagement would likely need to be provided and systematically faded in these cases (e.g., Potter et al., 2013). We also did not assess preference for the particular leisure or vocational tasks that were provided for Mark and Molly. It is possible that preference for the alternative tasks provided

during the S– component may moderate rates of functional item engagement. Future researchers may wish to investigate the use of chained schedules with participants who do not demonstrate any item engagement before intervention and with tasks of varying preference.

A second limitation is that although the chained schedule established inhibitory stimulus control of stereotypy, adult presence was still required to deliver tokens for appropriate item engagement and to implement the component changeover. An important area of future research may therefore be the elimination of adult-delivered tokens in the chained schedule. For example, it would be beneficial to investigate procedures to teach participants to deliver their own tokens after appropriate item engagement, to implement the component changeover after the required number of tokens are earned, and then to return to work after the reinforcement interval; other procedures might involve automated delivery of tokens (e.g., during computer work).

The type of correlated stimuli used to signal the S– and S+ components is another important and related area of future study. The correlated stimuli we used to establish stimulus control of stereotypy in the chained schedule were arbitrary rather than contextual. This may present a problem in some situations if the participant completes the necessary number of responses in the S– component and changeover to the S+ component occurs, but some other contextual feature of the situation indicates that stereotypy is actually not appropriate at the moment (e.g., when an adult delivers important instructions). Future research could evaluate several alternatives to the use of arbitrary correlated stimuli. One alternative could be the transfer of control from arbitrary to contextually relevant cues, or perhaps initiation of chained-schedule treatments using varied but contextually relevant cues. Another alternative could be a permission-based model in which participants are taught to mand for access to stereotypy, to tolerate the occasional denial of this mand and to engage in appropriate and contextually relevant responses when the mand for stereotypy is denied, and to engage in stereotypy only when a caregiver who has assessed the relevant contextual cues indicates that stereotypy is acceptable (e.g., similar to the treatment in Hanley et al., 2014, designed for socially mediated problem behavior).

Despite these limitations, which can be addressed in future research, we believe those who attempt to treat stereotypy should consider the use of chained schedules rather than single-component treatments (e.g., response blocking, enriched environment) or multiple schedules as a treatment for stereotypy because chained schedules may be more effective in reducing stereotypy, increasing item engagement, and establishing stimulus control, and may be preferred by those who experience the intervention.

REFERENCES

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

Ahrens, E. N., Lerman, D. C., Kodak, T., Worsdell, A. S., & Keegan, C. (2011). Further evaluation of response interruption and redirection as treatment for stereotypy. *Journal of Applied Behavior Analysis, 44*, 95–108. doi: 10.1901/jaba.2011.44-95

Bannerman, D. J., Sheldon, J. B., Sherman, J. A., & Harchik, A. E. (1990). Balancing the right to habilitation with the right to personal liberties: The rights of people with developmental disabilities to eat too many doughnuts and take a nap. *Journal of Applied Behavior Analysis, 23*, 79–89. doi:10.1901/jaba.1990.23-79

Bejarano, R., & Hackenberg, T. D. (2007). IRT-stimulus contingencies in chained schedules: Implications for the concept of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 88*, 215–227. doi:10.1901/jeab.2007.87-03

Brusa, E., & Richman, D. (2008). Developing stimulus control for occurrences of stereotypy exhibited by a child with autism. *International Journal of Behavioral Consultation and Therapy, 4*, 264–269. doi:10.1037/h0100855

Carroll, R. A., & Kodak, T. (2014). An evaluation of interrupted and uninterrupted measurement of vocal stereotypy on perceived treatment outcomes. *Journal of Applied Behavior Analysis, 47*, 264–276. doi:10.1002/jaba.118

Cassella, M. D., Sidener, T. M., Sidener, D. W., & Progar, P. R. (2011). Response interruption and redirection for vocal stereotypy in children with autism: A systematic replication. *Journal of Applied Behavior Analysis, 44*, 169–173. doi:10.1901/jaba.2011.44-169

Charlop, M. H., Kurtz, P. F., & Casey, F. G. (1990). Using aberrant behaviors as reinforcers for autistic children. *Journal of Applied Behavior Analysis, 23*, 163–181. doi:10.1901/jaba.1990.23-163

Conroy, M. A., Asmus, J. M., Sellers, J. A., & Ladwig, C. N. (2005). The use of an antecedent-based intervention to decrease stereotypic behavior in a general education classroom: A case study. *Focus on Autism and Other Developmental Disabilities, 20*, 223–230. doi:10.1177/10883576050200040401

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*, 106–124. doi:10.1002/bin.1378

Cunningham, A. B., & Schreibman, L. (2008). Stereotypy in autism: The importance of function. *Research in Autism Spectrum Disorders, 2*, 469–479. doi:10.1016/j.rasd.2007.09.006

Doughty, S. S., Anderson, C. M., Doughty, A. H., Williams, D. C., & Saunders, K. J. (2007). Discriminative control of punished stereotyped behavior in humans. *Journal of the Experimental Analysis of Behavior, 87*, 325–336. doi:10.1901/jeab.2007.39-05

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

Foster, L. G. (1998). Nervous habits and stereotyped behaviors in preschool children. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*, 711–717. doi:10.1097/00004583-199807000-00010

Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior, 91*, 257–286. doi:10.1901/jeab.2009.91-257

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

Haley, J. L., Heick, P. F., & Luiselli, J. K. (2010). Use of an antecedent intervention to decrease vocal stereotypy of a student with autism in the general education classroom. *Child & Family Behavior Therapy, 32*, 311–321. doi:10.1080/07317107.2010.515527

Hanley, G. P., Iwata, B. A., Thompson, R. H., & Lindberg, J. S. (2000). A component analysis of "stereotypy as reinforcement" for alternative behavior. *Journal of Applied Behavior Analysis, 33*, 285–297. doi:10.1901/jaba.2000.33-285

Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis, 47*, 16–36. doi:10.1002/jaba.106

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis, 30*, 459–473. doi:10.1901/jaba.1997.30-459

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)

Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 5*, 543–597. doi:10.1901/jeab.1962.5-s543

Kliebert, M. L., Tiger, J. H., & Toussaint, K. A. (2011). An approach to identifying the conditions under which response interruption will reduce automatically reinforced problem behavior. *Behavior Analysis in Practice, 4*, 17–26.

Koegel, R. L., & Covert, A. (1972). The relationship of self-stimulation to learning in autistic children. *Journal of Applied Behavior Analysis, 5*, 381–387. doi:10.1901/jaba.1972.5-381

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*, 1234–1243. doi:10.1016/j.rasd.2013.07.009

Leekam, S., Tandos, J., McConachie, H., Meins, E., Parkinson, K., Wright, C., … Le Couteur, A. (2007). Repetitive behaviours in typically developing 2-year-olds. *Journal of Child Psychology and Psychiatry, 48*, 1131–1138. doi:10.1111/j.1469-7610.2007.01778.x

Lovaas, O. I., Litrownik, A., & Mann, R. (1971). Response latencies to auditory stimuli in autistic children engaged in self-stimulatory behavior. *Behaviour Research and Therapy, 9*, 39–49. doi:10.1016/0005-7967(71)90035-0

Luczynski, K. C., & Hanley, G. P. (2009). Do children prefer contingencies? An evaluation of the efficacy of and preference for contingent versus noncontingent social reinforcement during play. *Journal of Applied Behavior Analysis, 42*, 511–525. doi:10.1901/jaba.2009.42-511

Luczynski, K. C., & Hanley, G. P. (2010). Examining the generality of children's preference for contingent reinforcement via extension to different responses, reinforcers, and schedules. *Journal of Applied Behavior*

Analysis, 43, 397–409. doi:10.1901/jaba.2010.43-397

Luczynski, K. C., & Hanley, G. P. (2014). How should periods without social interaction be scheduled? Children's preference for practical schedules of positive reinforcement. *Journal of Applied Behavior Analysis, 47*, 500–522. doi:10.1002/jaba.140

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

O'Connor, A. S., Prieto, J., Hoffmann, B., DeQuinzio, J. A., & Taylor, B. A. (2011). A stimulus control procedure to decrease motor and vocal stereotypy. *Behavioral Interventions, 26*, 231–242. doi:10.1002/bin.335

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. doi:10.1901/jaba.2000.33-13

Piazza, C. C., Hanley, G. P., & Fisher, W. W. (1996). Functional analysis and treatment of cigarette pica. *Journal of Applied Behavior Analysis, 29*, 437–450. doi:10.1901/jaba.1996.29-437

Potter, J. N., Hanley, G. P., Augustine, M., Clay, C. J., & Phelps, M. C. (2013). Treating stereotypy in adolescents diagnosed with autism by refining the tactic of "using stereotypy as reinforcement." *Journal of Applied Behavior Analysis, 46*, 407–423. doi:10.1002/jaba.52

Rafaeli-Mor, N., Foster, L., & Berkson, G. (1999). Self-reported body-rocking and other habits in college students. *American Journal on Mental Retardation, 104*, 1–10. doi:10.1352/0895-8017(1999)104<0001:SBAOHI>2.0.CO;2

Rapp, J. T., Patel, M. R., Ghezzi, P. M., O'Flaherty, C. H., & Titterington, C. J. (2009). Establishing stimulus control of vocal stereotypy displayed by young children with autism. *Behavioral Interventions, 24*, 85–105. doi:10.1002/bin.276

Rapp, J. T., Swanson, G., Sheridan, S. M., Enloe, K. A., Maltese, D., Sennott, L. A., … 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*, 527–547. doi:10.1016/j.ridd.2004.11.005

Tröster, H. (1994). Prevalence and functions of stereotyped behaviors in nonhandicapped children in residential care. *Journal of Abnormal Child Psychology, 22*, 79–97. doi:10.1007/BF02169257

Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis, 21*, 381–384. doi:10.1901/jaba.1988.21-381

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

Wunderlich, K. L., & Vollmer, T. R. (2015). Data analysis of response interruption and redirection as a treatment for vocal stereotypy. *Journal of Applied Behavior Analysis, 48*, 749–764. doi:10.1002/jaba.227

Received September 15, 2015
Final acceptance June 20, 2016
Action Editor, Joel Ringdahl