Special Issue: School-Based Interventions for Students with Autism Spectrum Disorder

Guest Editors: Cynthia M. Anderson, Tristram Smith, and Susan Wilczynski
Increasing Functional Leisure Engagement for Children With Autism Using Backward Chaining

Chelsea K. Edwards¹, Robin K. Landa¹,², Sarah E. Frampton¹, and M. Alice Shillingsburg¹,³

Abstract
Research with individuals with disabilities has demonstrated the utility of intervention approaches to address toy play, also referred to as functional leisure engagement (FLE). Examples include prompting FLE, blocking stereotypy, and differentially reinforcing appropriate FLE with social or automatic (i.e., access to stereotypy) reinforcers. Backward chaining has yet to be evaluated, but may be useful for establishing more complex FLE. The current study employed a treatment package consisting of these components with three school-aged children with autism in a therapeutic classroom. Effects were evaluated during pretest and posttest sessions, which consisted of free access to toys in a novel setting. The percentage of session with FLE was evaluated using a multiple probe design across participants. Results showed all participants demonstrated an increase in FLE and two participants showed decreased stereotypy. Feasibility for classroom implementation is discussed.

Keywords
autism spectrum disorder, backward chaining, functional engagement, functional play, toy play

¹Marcus Autism Center, Atlanta, GA, USA
²Western New England University, Springfield, MA, USA
³Emory University, Atlanta, GA, USA

Corresponding Author:
M. Alice Shillingsburg, Marcus Autism Center, 1920 Briarcliff Road NE, Atlanta, GA 30329, USA. Email: alice.shillingsburg@choa.org
The development of play skills is crucial for children diagnosed with autism spectrum disorder (ASD), as it functions as a platform for the development of social interaction and social communication skills (Bloom, 1993; Piaget, 1952). Although exploratory engagement with items is seen in infancy, functional toy play typically begins to emerge later in the first year of life. This higher level form of play has been delineated from its earliest precursors by being characterized as playing with toys “according to their intended function” (Baron-Cohen, 1987, p. 142). Demonstrations of functional play skills earlier in life have been correlated with the later development of language with typically developing children (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Ungerer & Sigman, 1984) and children with ASD (Sigman & Ruskin, 1999; Toth, Munson, Meltzoff, & Dawson, 2006). However, children with ASD engage in fewer spontaneous appropriate play acts than their typically developing peers (Lewis & Boucher, 1988; Wing, Gould, Yeates, & Brierley, 1977). Sigafoos, Roberts-Pennell, and Graves (1999) found that play skills for preschoolers with ASD showed little growth over a 3-year period in comparison with the development of their other adaptive skills. This deficit in play is of significant concern, as these failures to play may then exacerbate failures to develop social communication skills. Thus, interventions that support the development of play skills are of critical importance for learners with ASD (Jarrold, Boucher, & Smith, 1996).

Although a wide variety of interventions have been proposed to teach toy play to children with ASD (see Lang et al., 2009; Luckett, Bundy, & Roberts, 2007; Stahmer, Ingersoll, & Carter, 2003), many, especially highly structured, behavior-analytic procedures, have been described as failing to produce genuine play, in part due to the reliance on external rather than “internal” rewards and limited spontaneity, flexibility, and generalization of targeted play behaviors (Luckett et al., 2007). This criticism highlights several key features of play that researchers and clinicians have emphasized including that play is a behavior that is “freely chosen, personally directed, and intrinsically motivated” (Hughes, 2012). Thus, instead of being highly structured and adult directed, activities should be selected and initiated by the child. Only interventions that evaluate play in an unstructured context, in which the individual is free to engage with the toys or in other behaviors, would meet this criteria (e.g., Eason, White, & Newsom, 1982; Nuzzolo-Gomez, Leonard, Ortiz, Rivera, & Greer, 2002). Furthermore, play should not be motivated exclusively by contingent reinforcers such as adult attention or other tangible reinforcers. Thus, the intervention, including any contingent reinforcers, must be fully withdrawn for this criteria to be met (e.g., Eason et al., 1982; Kasari, Gulsrud, Wong, Kwon, & Locke, 2010; Nuzzolo-Gomez et al., 2002). These characteristics of an intervention not only increase the acceptability of
outcomes, but also make it more sustainable in applied settings. For example, a teacher serving 6 to 10 children with ASD will find it difficult to deliver carefully scheduled tangible reinforcers during common free play contexts, in the midst of other responsibilities.

Another key feature of play to consider is the importance of generalization. Play outcomes should also have generality across toys and settings. Without these measures, there is little to suggest that the play intervention will have effects beyond the intervention setting itself. This is crucial when services are delivered outside of the child’s home (i.e., in a treatment clinic or school), because play is an ongoing, daily activity for children. Last, in addition to the quantity of play behaviors, the quality of play should be emphasized. Several studies have included dependent measures such as simple object manipulation (Hanley, Iwata, Thompson, & Lindberg, 2000) and touching toys (Nuzzolo-Gomez et al., 2002), which may have been appropriate for the participants in those studies but limits extensions to more complex play behavior requiring a series of steps.

In addition to these key elements of what constitutes play, another factor to consider when designing a toy play intervention in an applied setting is whether the intervention will address sources of competing reinforcers. Engagement in other behaviors that produce competing reinforcers have been cited as a potential reason that children with ASD fail to develop appropriate play skills, most notably when the competing behavior is stereotypy (e.g., Honey, Leekam, Turner, & McConachie, 2006; Wing & Gould, 1979). Thus, many studies evaluating treatment packages for addressing toy play include measures to specifically target the reduction of stereotypy. Some components of these treatment packages include contingent access to tangible and social reinforcement for appropriate play (e.g., Eason et al., 1982; Horner, 1980; Nuzzolo-Gomez et al., 2002; Singh & Millichamp, 1987), blocking stereotypy (e.g., Hanley et al., 2000), and contingent access to stereotypy for appropriate play (e.g., Hanley et al., 2000; Potter, Hanley, Augustine, Clay, & Phelps, 2013). Thus, not only will an effective treatment package need to address the common criticisms of behavioral play interventions noted above, it must also effectively address stereotypy.

Lang et al. (2014) provided an efficacious treatment package that met these criteria with three young children with ASD who demonstrated low levels of play and frequent stereotypy. To address concerns related to generality, the baseline and posttest measures were conducted in a different setting from intervention. Furthermore, generalization to untrained toy sets was also measured. To address concerns related to extrinsic reinforcers, the baseline and posttest evaluations took place during unstructured play situations in the absence of the teacher who conducted the treatment. Treatment consisted of
redirecting stereotypy, prompting toy play, and contingent tangible reinforcement on a gradually faded schedule. All participants showed increased occurrences of appropriate play in the untrained setting, in the absence of the teacher and extrinsic reinforcement. For one participant, play behaviors generalized to an untrained play set and stereotypy remained at low levels. The other two participants still showed a moderate degree of stereotypy, though below baseline levels, and generalization across toy sets was not observed. For these participants, a lag schedule of reinforcement was added which resulted in sustained, low levels of stereotypy and generalization to the untrained play set. Overall, this study serves as a strong example of a treatment package to address play skills in an applied, school-like setting. However, the children in this study were 3 years old and had emerging vocal language so their responsiveness to this intervention may not represent findings for older, school-aged students with greater barriers to learning. Furthermore, participants selected for the Lang study engaged in pretreatment repetitive toy play and were able to select toys within a paired stimulus assessment, suggesting they already had some degree of interest in engaging with toys. Two other studies also showed both a reduction in stereotypy and increased play with generality across toys, one using video modeling (Paterson & Arco, 2007) and the other, self-management (Stahmer & Schriebman, 1992). While the participants in these studies were school-aged (6-13 years), they were also able to communicate using complete sentences and were reported to have generalized imitation repertoires (Paterson & Arco). So, although these interventions addressed stereotypy and generalization of play behaviors, these approaches may not be effective for minimally verbal, school-aged children with ASD.

Although several effective approaches to increasing toy play while also reducing stereotypy have been established in the literature, no study has shown how to teach more complex play sequences to minimally verbal, school-aged children. This may be accomplished through the application of backward chaining, a procedure often referenced in curricular manuals (e.g., Leaf & McEachin, 1999; Maurice, Green, & Luce, 1996) but yet to be widely applied in the literature. In their review of the literature, Stahmer et al. (2003) outlined general recommendations for teaching simple toy play to individuals with disabilities using chaining procedures. The authors highlighted the importance of breaking down a toy play sequence into steps, teaching the individual steps, and then applying forward or backward chaining procedures to ensure the sequence of steps is established so that the toy play is functional. Given the complexity of play skills and the likelihood that functional play will require multistep sequences, backward chaining may be a valuable procedure to include in treatment packages aimed at increasing play.
The purpose of the current study was to build upon prior research on establishing functional play skills by evaluating the direct and indirect effects of an intervention package consisting of (a) backward chaining, (b) blocking stereotypy, and (c) differential reinforcement in the form of tangible items and access to stereotypy. Furthermore, the current study addressed limitations of prior studies by including measures of play during free play, assessment of generalization across settings, and generalization to novel toys. These components were included to most closely mimic those present in most classroom settings. Importantly, the intervention was conducted with three school-aged children with ASD with minimal verbal skills to evaluate the effectiveness of the intervention within a population with more pronounced deficits than has been included in prior studies (e.g., Lang et al., 2014; Nuzzolo-Gomez et al., 2002; Paterson & Arco, 2007; Stahmer & Schriebman, 1992).

Method

Participants and Setting

All participants received one-on-one services within a therapeutic classroom at a center-based intervention program, 5 days a week for 2 hr each day. The Verbal Behavior-Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) was administered to each participant upon their admission into the clinic and was used for ongoing clinical programming. The clinical case manager, a board certified behavior analyst (BCBA), in consultation with the participant’s parents determined that an intervention to support the development of functional play was appropriate. Scores from the VB-MAPP were used to confirm deficits with play, but no other inclusion or exclusion criteria were required for participation in the study.

Ethan, a 4-year-old male, was diagnosed with ASD at the age of 3 by a licensed psychologist. He was administered the Mullen Scales of Early Learning (Mullen, 1995), which indicated his skills were in the 18- to 21-month range for receptive language, visual reception, and fine motor skills while scores related to oral production were comparable to 14 months. The Adaptive Behavior Assessment System (ABAS; Harrison & Oakland, 2003) indicated Ethan’s overall functioning was below the 1st percentile. Ethan had a VB-MAPP Milestone score of 46, a play domain score of 7.5 (18- to 30-month range), and an imitation domain score of 7 (18- to 30-month range). Specifically, in the area of play, Ethan was observed to engage functionally with only a few items when specifically directed, but was unable to assemble toys with multiple parts. Although Ethan engaged in cooperative
mouthing, disruption, etc.), or engaging with the toy for less than 5 s at a time. Further, FLE did not require play responses to occur in a specific order or for completion of a sequence of actions. The duration of FLE was tracked with a 5-s onset/offset. The number of seconds spent in FLE was then divided by 360 (the total number of seconds in the 5-min session), resulting in the percentage of session spent engaged in FLE.

Step completion served as the dependent variable during step completion baseline and treatment probe sessions. Step completion was defined as the independent completion of a step in the designated order. A task analysis was created for each activity consisting of eight steps. Step completion was defined separately for each toy and was scored each time the participant completed a step in the designated order, without prompts. Step completion was not recorded if the participant did not engage with the toy or complete the step in order, or required prompts to complete the step. For example, for the train set containing six tracks and one train, 100% correct step completion was recorded if the participant completed the following steps in order: Step 1–connect track, Step 2–connect track, Step 3–connect track, Step 4–connect track, Step 5–connect track, Step 6–connect track, Step 7–put train on top of track, Step 8–move train along the track. Observers recorded the total number of steps completed, which was divided by the total number of steps and converted to a percentage. As backward chaining was used, the chain was completed by the therapist and only the targeted steps were probed.

Stereotypy was measured during pre- and posttests for Trey and Lance and included a variety of behaviors. For Trey, stereotypy was defined as any instance of mouthing, spinning, posturing, and bursts of running. Mouthing was defined as any instance in which an object or Trey’s fingers crossed the plane of Trey’s lips with a 5-s onset/offset. Spinning was defined as any instance in which Trey rotated his body 360 degrees with a 5-s onset/offset. Posturing was defined as any instance in which Trey’s knees were on the floor with a 5-s onset/offset. Bursts of running were defined as any instance in which Trey moved at a pace faster than walking around the room with a 5-s onset/offset. For Lance, stereotypy was defined as any instance of head hitting, pacing, repetitive body movements, scratching, and spinning. Head hitting was defined as any instance or attempt in which an object Lance was holding contacted his head from a distance of 3 in. (7.6 cm) or greater. Pacing was defined as any instance in which Lance walked more than 3 ft. (.91 m) from one point in the room to another, at least two times with a 5-s onset/offset. A repetitive body movement was defined as any instance in which Lance engaged in behavior at least five consecutive times with no functional relation to the toys in the room with a 5-s onset/offset. Scratching was defined as any instance in which Lance’s nails came into contact with his skin and left
red marks with a 5-s onset/offset. Spinning was defined as any instance in which Lance rotated his body 360 degrees with a 5-s onset/offset. A total duration measure was taken for all behaviors and converted into a percentage of session.

**Interobserver Agreement (IOA)**

IOA data were collected by trained independent observers using pen and paper and were collected for FLE during an average of 79.8% (range = 66.8%-100%) of sessions across all participants. IOA data were collected for stereotypy an average of 34.8% (range = 33.3%-36.4%) of sessions across two participants (Lance and Trey). The percentage of IOA data was calculated by taking the smallest duration measure recorded and dividing by the largest duration measure, then multiplying by 100. The percentage of IOA averaged 96.7% (range = 90%-100%) for FLE. The percentage of IOA averaged 87.1% (range = 72.8%-100%) for stereotypy. For step completion data, data were collected for an average of 44.5% (range = 41%-50%) of sessions across all participants. For these data, trial-by-trial IOA was calculated for each participant by taking the number of trials with agreements and dividing by the total number of trials with agreements and disagreements then multiplying by 100. The percentage of agreements averaged 98.9% (ranging = 96.7%-100%).

**Fidelity of Implementation**

Treatment integrity data were also collected by a trained independent observer. Step-by-step procedures were outlined in a checklist, and the independent observer recorded whether the therapist followed the procedures exactly as detailed (scored as a “+”) or with any deviation (scored as a “−”). Scores were calculated by dividing the number of correctly implemented steps by the total number of steps then multiplying by 100. The average percentage of sessions in which treatment integrity was scored across all participants was 61% (range = 44.5%-79.8%). The percentage of correct implementation averaged 100%.

**Experimental Design and Visual Display of Data**

A multiple probe design across participants was used to evaluate FLE prior to and following step completion treatment (i.e., pretest and posttest). The pretest and posttest sessions followed the procedures of a free operant preference assessment (Roane, Vollmer, Ringdahl, & Marcus, 1998). Step completion
was evaluated in a manner similar to a changing criterion design during baseline and treatment probe sessions. However, as backward chaining was incorporated into the treatment probe sessions, the participants were not free to complete all of the possible steps. Thus, the number of steps that could be completed in a treatment probe was limited to the target step and previously mastered steps.

**Procedures**

*Reinforcer selection.* Therapists identified preferred items for all phases of this study using a variety of formats. Because the participants in this study were also being served in a clinic providing comprehensive intervention, careful assessment of highly preferred items and activities using questionnaires (Fisher, Piazza, Bowman, & Amari, 1996) and structured preference assessments (e.g., DeLeon & Iwata, 1996) were periodically conducted. The identified items and activities were then incorporated into sessions and moment to moment verbal requests of those identified items were used to identify current preference. Therefore, a variety of preferred items identified via parental interview, structured preference assessment, and requests were present on the table and additional items were stored in bookshelves nearby. The therapist either waited for participants to engage in a request or issued a vocal prompt to occasion a request (e.g., “What would you like to work for?”) before starting each session. Lance typically requested the iPad, cookies, and pretzels. For Trey, an iPad was the only item requested as he did not exhibit sustained interest in any other play activities or foods. Ethan typically requested potato chips, candy, or the iPad. Ethan’s clinical programming also included the use of a token board. Once all tokens were earned, he would often clap his hands and remove the tokens to start over, rather than exchange them for a backup reinforcer. Thus, the token board was incorporated throughout all phases of the study for Ethan. Tokens (Ethan) and other items (Lance and Trey) were delivered at the end of each session. During these reinforcer intervals, therapists reinforced requests (including bids for attention) and allowed unrestricted access to stereotypy.

*Pretest and posttest.* Pretest and posttest sessions were used to evaluate the percentage of session in which the participant demonstrated FLE and stereotypy. Sessions were 5 min and were conducted in a separate room inside the clinic. Procedures followed those used in free operant preference assessments (Roane et al., 1998). Toys were placed on a table in the room and were equidistant from each other and the participant. Using hand over hand prompts provided by the therapist, the participant was exposed to all toys individually.
for 10 s prior to every session. The therapist began the session by stating, “Play with whatever you want.” During the session, no prompts or other forms of interaction with the participant were provided. If dangerous behavior occurred, such as elopement, this behavior was blocked, and minimal attention was provided. The therapist tracked the total duration of FLE and stereotypy on separate timers and a count of the number of toys the participant demonstrated FLE with during the session. After 5 min, the therapist provided noncontingent access to tokens (for Ethan) or the previously identified preferred item (Lance and Trey) for a period of 1 min. Noncontingent access was provided so that results of pre- and posttest sessions could be evaluated in the absence of programmed contingencies. Furthermore, it was hypothesized that these conditions would mimic those of a school routine. For example, after the independent play time, all students rotate to recess or snack (presumed highly preferred activities) regardless of performance during play time. Additionally, the participants were not blocked from engaging in stereotypy during these periods. Pretest sessions were conducted until stability.

**Step completion baseline.** Baseline sessions were conducted with each of the six toys. Therapists presented one toy per session, in random order, for 2 min, during which data collectors scored the number of steps completed. During the session, no prompts or other forms of interaction with the participant were provided. Dangerous behavior, such as elopement, was blocked, but the therapist avoided redirecting the participant to engage in any other behavior. At the end of the session, therapists delivered tokens (Ethan, which he then exchanged for his iPad) or noncontingent access to the previously selected preferred items for 2 min. As noted above, during these periods of noncontingent access, stereotypy was not blocked. After the period of noncontingent access, the therapist presented the next session. A minimum of three baseline sessions per toy were conducted until data were deemed stable.

**Step completion treatment probe.** Percentage of steps completed for each of the three toys assigned to treatment was evaluated during each step completion treatment probe session. The therapist randomized the toys and placed them one at a time in front of the participant. Using backward chaining, the therapist completed all nontargeted steps then provided the instruction, “Play with your toys” and provided an independent response opportunity to complete the next step. The therapist then recorded the steps completed independently by the participant. If the participant errored or did not initiate step completion within 10 s, the therapist used least-to-most prompting to complete that step and then moved on to the next step to be probed. Least-to-most
prompting progressed through the following as needed: a verbal instruction followed by a gestural prompt, followed by light physical guidance to move the child’s hands to the toy, followed by hand-over-hand guidance to complete the step. Upon independent completion of the behavior chain, the therapist provided 30-s to 60-s access to reinforcers depending on the number of steps completed. Step completion treatment probe sessions were then followed by treatment sessions (described below). Mastery criteria ranged from three to six consecutive probes with 100% independence on the targeted step. Following mastery of a step, treatment began for the next step in the chain. Mastery of the entire chain for each toy was 100% independent step completion of a predetermined number of consecutive probes; however, probes of all steps continued following mastery to assess maintenance.

**Step completion treatment.** Following step completion treatment probes, step completion treatment sessions were conducted with each toy. Procedures were similar to step completion treatment probes; however, the therapist used most-to-least prompting to help the participant to complete the target steps and provided the preselected preferred items (e.g., tokens, tangible items) contingent upon completion of the final step of the chain. If stereotypy occurred during the treatment sessions, the therapist immediately blocked and redirected by prompting the appropriate step in the chain. Each step completion treatment sessions typically lasted less than 2 min. Following three step completion treatment sessions, a step completion treatment probe session was conducted. Of note, a treatment probe was not conducted until at least 20 min had elapsed since the last step completion treatment session.

**Results**

Results of pretests and posttest for all participants are shown in Figure 1 and the results of step completion baseline and treatment probes are shown for individual participants in Figures 2 to 4. The average percentage of sessions (pretest and posttest) with occurrences of stereotypy for Lance and Trey is shown in Figure 5. As shown in Figure 1, Ethan demonstrated no FLE during pretest sessions \( (M = 0\%) \). In Figure 2, the step completion baseline data show he completed an average of 16.7% steps correctly for piano \( (\text{range} = 0\%-25\%) \), an average of 29.2% steps correctly for puzzles \( (\text{range} = 0\%-87.5\%) \), and no steps correctly for the trains. During step completion treatment, Ethan reached mastery criteria, independently completing 100% of steps correctly for at least three consecutive sessions, in 36, 34, and 46 sessions for piano, puzzles, and trains, respectively. During the posttests (Figure 1), Ethan engaged in FLE for an average of 66.7% of the sessions
Figure 1. Results from pretest and posttest sessions are displayed for all participants.

Note. The percentage of sessions in which the participants engaged in FLE with target toys are shown with black bars. The percentage of sessions in which participants engaged in FLE with generalization toys are shown with gray bars. The asterisks indicate that a session was conducted, but 0% FLE was exhibited. The solid phase line indicates when SC treatment occurred with toys assigned to treatment. SC = step completion; FLE = functional leisure engagement.
average duration of 35.5% of the sessions (range = 20%-46%). In Figure 3, the step completion baseline data show he completed 0% of steps correctly for piano, puzzles, and trains. During step completion treatment, Lance reached mastery criteria in 41, 34, and 45 sessions for piano, puzzles, and trains.
trains, respectively. During the posttests (Figure 1), Lance engaged in FLE for an average of 75.2% of the sessions (range = 72%-79.3%). The percentage of session engaged in FLE with generalization toys averaged 23.1% (range = 11.7%-32.7%) and for treatment toys averaged 52.1% (range = 39.3%-67.7%). As shown in Figure 5, during posttests Lance engaged in stereotypy for an average duration of 2.2% of the sessions (range = 0%-6.6%).

As shown in Figure 1, Trey displayed FLE during the pretests for an average of 5.9% of the sessions (range = 0%-29.3%). Trey engaged only with the treatment toys during the pretest. As shown in Figure 5, during pretests Trey engaged in stereotypy for an average duration of 40.7% of sessions (range = 3.6%-65%). As shown in Figure 4, the step completion baseline data show he completed an average of 0% steps correctly for piano, an average of 16.67%
of steps correctly for puzzles (range = 12.5%-25%), and an average of 33.3% of steps correctly for trains (range = 0%-62.5%). During step completion treatment, Trey reached mastery criteria in 94, 36, and 71 sessions for piano, puzzles, and trains, respectively. During the posttests (Figure 1), Trey engaged in FLE for an average of 52.11% of the sessions (range = 0%-69.7%). The percentage of session engaged in FLE with generalization toys averaged 14.7% (range = 0%-23%) and for treatment toys averaged 37.4% (range = 0%-55.4%). As shown in Figure 5, during posttests Trey engaged in stereotypy for an average duration of 17.6% of the sessions (range = 0%-38.4%).

**Discussion**

All three participants in the current study demonstrated an increase in functional play in treatment settings (i.e., step completion treatment), which resulted in an increase in FLE in a free play setting (i.e., posttest). Two of the three participants engaged with both targeted and generalization toys in the posttests. The results of this study extend upon prior research on teaching functional toy play to children with autism by demonstrating the efficacy of backward chaining in teaching more complex play behaviors involving multiple steps in a sequence. These results also add to the body of research that demonstrates the generality of behavior-analytic interventions for teaching toy play and address some of the limitations regarding behavioral approaches to teaching play. By assessing toy play in an unstructured setting in which the intervention was fully withdrawn, we were able to evaluate whether activities were selected and initiated by the child, who was free to either play with items or engage in other behaviors (e.g., stereotypy, impassivity, or attempts to gain access to other reinforcers). The inclusion of generalization toys enabled us to assess treatment generality to novel items and allowed us to identify the occurrence of novel play behaviors. Our results indicated that, following treatment, participants continued to engage in play in the absence of programmed contingencies, thereby meeting the definition of play set forth by Hughes (2012) and addressing the concern proposed by Luckett et al. (2007) that highly structured approaches result in reliance on external rewards. The increase in untargeted play actions with novel toys is also notable given Luckett et al.’s concerns regarding limited spontaneity, flexibility, and generalization associated with highly structured behavior-analytic interventions. Although similar effects have been previously demonstrated (Lang et al., 2014), our results provide further support to the generality of behavioral interventions for play by including children who exhibited limited interest in toys prior to intervention (i.e., children who either did not approach toys or did not engage in sustained interaction with toys) and with more
pronounced intellectual and verbal deficits than in other studies (Paterson & Arco, 2007; Stahmer & Schriebman, 1992).

The generality of play behaviors to untargeted toys and the persistence of play even when external reinforcers were removed suggests that something about the learned play behaviors themselves were reinforcing. This is an ideal outcome. The use of a backward chaining procedure to teach multistep functional play potentially lead to the child experiencing for the first time, the reinforcing experience of playing with toys. In contrast, previous studies that targeted simple object manipulation or merely touching toys may not have taught behaviors that ultimately were automatically reinforcing. In the current study, it is possible that teaching the participants what to do with the toys required reinforcement for the behaviors to be learned, but did not require external reinforcement for the behaviors to be maintained. Further, the generalization toys may have had similar features to the trained toys that made engaging in a sequence of responding possible with those toys even though they were novel. For example, the car and car track have similar features to the train and train track in that the car and train both have wheels and are vehicles and the car track and train track are both flat tracks with grooves (though different shapes and colors). Thus, in the presence of the novel car track, similar responses such as pushing the car along the car track may have been evoked. Teachers and parents hoping to teach generalized toy play skills may want to consider whether the toy play behaviors will be automatically reinforcing once taught and may consider having other toys with similar features available.

The treatment package we employed involved several components of previously established interventions for increasing functional engagement such as prompting and differential reinforcement of appropriate play behaviors with social reinforcers (e.g., Nuzzolo-Gomez et al., 2002), restricted access to stereotypy during “play time” (e.g., Hanley et al., 2000; Potter et al., 2013), and access to nonsocial reinforcers contingent on appropriate play (i.e., contingent access to stereotypy or other behaviors; e.g., Potter et al., 2013), which were embedded within a backward chaining approach. Although this intervention sounds quite intensive, one backward chaining session with one targeted toy typically took fewer than 2 min to conduct. Therefore, while other interventions, such as self-management or in-vivo or video modeling, might seem relatively more feasible to implement in settings such as the classroom, the current intervention package is fairly simple to implement and could be interspersed with other teaching activities throughout a school day.

The intervention employed in our study was efficacious for three participants who exhibited limited to no play in unstructured settings, with limited imitation skills, and who engaged in stereotypic behaviors (Lance and Trey), suggesting that this package may be particularly suited for children with
more pronounced deficits and without the prerequisites (i.e., imitation skills) to benefit from other empirically supported interventions such as video modeling or self-management. In addition, the finding that it resulted in generalization of play to contexts with minimal to no oversight suggests that, once the toy play skills are taught, parents or teachers may see increases in play during periods in which attention is not continuously available, without having to intervene in or otherwise restructure the typical play context. This may be ideal for caregivers whose attention must be diverted during the times in which the child is expected to play independently. In a classroom, for example, free play contexts are often associated with lower instructor to child ratios as the teacher is preparing for the next instructional block.

As noted, two of the participants in our study demonstrated reductions in stereotypy in free play contexts following intervention. Given the likelihood that stereotypy competes with appropriate functional play, this result is encouraging. However, it is important to note that these participants engaged in stereotypy at low to moderate levels in baseline. Whether the treatment package employed in our study would also be successful for individuals who engage in high-rate or continuous stereotypy is unclear and other procedural components may be necessary to facilitate practical maintenance of low rate stereotypy for these individuals (see Slaton & Hanley, 2016, for an example and further discussion). Thus, replication across participants is necessary to evaluate this across a wide variety of children with ASD. Future researchers should also consider conducting component analyses to determine the critical components of the intervention package. Last, additional research is also needed to evaluate whether similar results would be achieved when the intervention package is implemented by teachers or parents and whether caregivers deem the procedures and outcomes associated with this intervention as socially acceptable.

A few limitations and other procedural aspects of our study indicate several other areas for future research. First, we assessed generalization across only one setting and did not evaluate maintenance in the free play, posttest context. Therefore, in addition to researching the generality of this intervention across participants, researchers should examine the extent to which this intervention facilitates both generalization and maintenance across a wider variety of situations. Second, we included a component (i.e., backward chaining) that has not been previously evaluated in the literature on toy play. Though effective in the current study, it is not clear if other procedures would be more effective and efficient. Therefore, future researchers might investigate the relative efficacy of backward chaining compared with other teaching procedures, such as forward chaining or differential reinforcement with schedule thinning, in promoting generalization and maintenance of toy play among children who demonstrate limited interest in toys. Furthermore,
it may be the case that the relative efficacy of backward chaining is dependent on the type of play skills being taught and, therefore, researchers should also consider assessing the relative efficacy of backward chaining using a wide variety of play items, including those that are commonly associated with a specific sequence of actions (e.g., a puzzle set) and those that are not (e.g., dolls). Third, we did not formally assess preference for the toys prior to selecting them for inclusion in the study. Given that the participants did not engage in any play prior to this study, selecting toys based on preference was not feasible. However, if a child shows interest in toys or engages in some approximations of play with toys, conducting a preference assessment and including higher preferred toys is warranted. Finally, we did not collect data on whether participants would have engaged in steps above and beyond that which we specifically reinforced during the backward chaining intervention. In other words, participants were given the opportunity to complete only a certain number of steps at the time of the treatment probe—whether they would have exceeded the target requirement is unclear, but is possible, given the similarity of some steps (e.g., putting one vs. another piece in a puzzle). Therefore, during treatment, future researchers and practitioners may also consider interspersing terminal probes that enable students to engage in multiple play actions, to determine whether play is conforming to the contingencies in place.

In summary, the current treatment package was effective in teaching functional toy play for three participants diagnosed with autism and may be useful for parents and teachers who are interested in promoting play. When considering this strategy for use in a school setting, it may be useful to select toys that promote developmentally appropriate, functional play and that result in a specific outcome when used correctly. Therefore, it is important that the selection of toys matches the developmental level of the child and, when feasible, takes into consideration the child’s preferences for the types of toys and activities used. It is also important to understand that external reinforcement may be needed to teach play behaviors and that once the natural reinforcement of the new play behaviors is experienced, those external reinforcers can be removed. Last, teachers may appreciate the current strategy in that each teaching session only required a few minutes making this procedure something that might easily fit in with other activities that occur during the school day. Though not explicitly evaluated, it may be easier to teach step completion in a few brief 1:1 sessions throughout the day.

Acknowledgments
The authors would like to thank Mandi Kaur and Heidi Morgan for their assistance with data collection.


**Author Biographies**

**Chelsea K. Edwards** received her BA in psychology at Agnes Scott College in Decatur, GA, and her MA in applied behavior analysis at Ball State University in Muncie, IN. She currently resides in Duluth, GA and is a practicing BCBA.

**Robin K. Landa**, MS, BCBA, is a doctoral student at Western New England University. Her current research interests include assessment and treatment of problem behavior and procedures for promoting generalization and maintenance of adaptive skills.
Sarah E. Frampton, MA, BCBA, is a behavioral clinical practitioner at the Marcus Autism Center. Her current research interests include procedures for promoting emergence of untrained skills and developing social communication skills for children with autism spectrum disorder.

M. Alice Shillingsburg is the director of the Language and Learning Clinic at the Marcus Autism Center. She is an assistant professor in the Department of Pediatrics at Emory University School of Medicine. She is a licensed psychologist and board certified behavior analyst.