5-2017

Evaluation of Multiple Exemplar Training Plus Discrimination Training on Promoting Generalization of Response Variability

Bethany P. Contreras
Utah State University

EVALUATION OF MULTIPLE EXEMPLAR TRAINING PLUS DISCRIMINATION TRAINING TO PROMOTE GENERALIZATION OF RESPONSE VARIABILITY

by

Bethany P. Contreras

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Disability Disciplines

Approved:

Thomas S. Higbee, Ph.D. Major Professor	Sarah E. Pinkelman, Ph.D. Committee Member
Timothy A. Slocum, Ph.D. Committee Member	Amy Odum, Ph.D. Committee Member
Tyra Sellers, Ph.D. Committee Member	Mark R. McLellan, Ph.D. Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2017

ABSTRACT

Evaluation of Multiple Exemplar Training Plus Discrimination Training On Promoting Generalization of Response Variability

by

Bethany P. Contreras, Doctor of Philosophy

Utah State University, 2017

Major Professor: Thomas S. Higbee, Ph.D.
Department: Special Education and Rehabilitation

One of the defining characteristics of autism is the presence of excessive repetitive behaviors. Many children with autism engage in rigid and repetitive play. Researchers have shown that variability of play behavior, among other behaviors, can be increased through contingencies of reinforcement. However, little is known regarding generalization of response variability beyond the specific responses that are trained. The purpose of this dissertation was to evaluate the effects of combining multiple exemplar training with discrimination training on increasing varied play behavior with trained play materials and on promoting generalization of varied play behavior to untrained play materials. We conducted multiple exemplar training plus discrimination training with two participants, but did not observe discriminated responding with either. After discontinuing the discrimination training component, we observed increases in varied responding under the multiple exemplar training alone for both participants (i.e., lag

schedules alone). For the third participant, we implemented multiple exemplar training alone and also observed increases in varied play behaviors during this phase. For all three participants, we observed generalization of varied play behavior to at least some untrained play sets.

(121 pages)

PUBLIC ABSTRACT

Evaluation of Multiple Exemplar Training Plus Discrimination Training On Promoting Generalization of Response Variability

Bethany P. Contreras

Typically developing children learn from play. For example, play serves as a foundation for children to acquire early language and social skills. Children with autism tend to have deficits in play, and often engage in rigid or repetitive behaviors during play. Such rigid play behavior can limit opportunities for these children to learn from play. Researchers have shown that it is possible to increase the variety of play behaviors that children with autism engage in. But, research has not yet shown whether these gains in play behavior will transfer to other play environments and situations. Therefore, the purpose of this study was to investigate methods for promoting the transfer of varied and appropriate play to other play situations with three children with autism. In this study, we increased varied play behavior by providing rewards for playing in a varied manner (and not providing rewards for playing in an inappropriate or rigid manner). We did this with multiple different play situations to help the participants learn to engage in varied play in different situations. We then tested to see if the participants would vary their play with completely new play situations. We found that, following some modifications, our procedures were successful at increasing varied play behavior for all three participants, and that their varied play transferred to other play situations.

ACKNOWLEDGMENTS

First, I want to thank my advisor, Thomas Higbee, for his continued support and guidance throughout the dissertation process and during my 4 years in this program. The road was not always easy or straight, but the journey was worth it. I also want to thank my committee members, Timothy Slocum, Tyra Sellers, Sarah Pinkelman, and Amy Odum for their thoughtful comments and support on this project.

To my lab mates, especially Annie, Azure, Lorraine, Victoria, Amy, and Stephanie, thank you for everything. Without your support and dedication, I would never have been able to complete this project. I cannot fully express how much I appreciate and am indebted to you for all of the time you spent running and coding sessions.

To Audrey Hoffmann and Jay Hinnenkamp, thank you for being my sounding board. Thank you for being my emotional support. Thank you for just being there.

To my husband, Kevin, thank you for your love and support. Thank you for encouraging me to apply to this program, thank you for being willing to move across the country, thank you for being the world's best dad while I was busy being a full-time grad student. To Oliver and Gus, thank you for being the best two kids anyone could ask for. Your smiles and hugs kept me going every day.

Bethany P. Contreras

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Play serves many important roles in a child's learning and development, such as providing a context for social interactions with peers (McConnell, 2002), increasing the child's ability to learn in natural settings (Buysse, Wesley, Keyes, & Bailey, 1996), and laying the foundation for developing leisure skills (Barton & Wolery, 2008). Development of play skills has also been linked to development of language (McCune, 1995). Children with autism often have severe deficits in play skills. Researchers have reported that the functional and symbolic play of individuals with autism is less elaborate and less varied than that of typically developing peers (Williams, Reddy, & Costall, 2001). Several researchers have also noted that individuals with autism lack the generative property of play (e.g., Baron-Cohen, 1987; Charman & Baron-Cohen, 1997; Jarrold, Boucher, & Smith, 1996). In other words, many individuals with autism can engage in a variety of play behaviors when prompted, but engage in less varied or novel play independently or spontaneously when compared to typically developing peers. These reports are not surprising, given that one of the defining characteristics of autism is the presence of excessive repetitive behavior (American Psychiatric Association [APA], 2013). As summarized by Rodriguez and Thompson (2015), invariant responding across a variety of response topographies is a general characteristic of individuals with autism spectrum disorder (ASD).

Invariant behavior may be maladaptive in some situations. For example, when you have lost your keys, engaging in invariant behavior by looking in the same place over

and over again is unlikely to result in reinforcement. Engaging in more varied behavior, say by looking in different places, is more likely to result in reinforcement (i.e., finding the keys). Response variability can be thought of as an adaptive feature of behavior, meaning that varied responding can result in access to reinforcement. However, response variability may not be adaptive in general, but rather in certain situations. That is, in certain situations high levels of response variability may result in reinforcement and in other situations varied responding might result in extinction or punishment. For example, varying the answer to the question, "what is your name?" would be maladaptive.

When it comes to play, children with autism often display invariant behavior that may be maladaptive and reduce the amount and type of reinforcement they access. For example, a child might engage in invariant play by selecting the same items to play with or repeating a small number of play actions over and over again. Low response variability may limit the amount of reinforcement received compared to what is available (Rodriguez & Thompson, 2015; Wolfe, Slocum, & Kunnavatana, 2014). For example, it may be reinforcing for a child to receive attention from peers or siblings during play. However, if the child continues to do the same play action over and over again, the peers may lose interest in playing and will leave to do something else. In this example, reinforcement in the form of attention was available, but was not accessed due to excessive repetitive behavior. Low response variability may also reduce the number of opportunities for social interaction and language development. For example, if a child who is participating in child-led language training (such as pivotal response training) is engaging in the same play actions repetitively, this will reduce the amount and type of

language that can be prompted and reinforced. Additionally, invariant behavior can be socially stigmatizing to the individual with ASD, which can further limit opportunities to build social skills. Therefore, it is important to promote variability of play behaviors in individuals with autism.

To date, several researchers have demonstrated that variability of a variety of behaviors emitted by individuals with ASD can be directly influenced through contingencies of reinforcement. Responses investigated include manding (e.g., Brodhead, Higbee, Gerencser, & Akers, 2016), intraverbal responding (e.g., Contreras & Betz, 2016), tacting (e.g., Heldt & Schlinger, 2012), echoic responding (e.g., Esch, Esch, & Love, 2009), button pushing (e.g., Miller & Neurginer, 2000), selecting activities (e.g., Cammilleri & Hanley, 2005), and play behaviors (e.g., Baruni, Rapp, Lipe, & Novotny, 2014). Although studies have demonstrated that response variability emitted by individuals with autism can be increased, it is largely unknown if these increases generalized beyond the specific behaviors that were targeted.

Stokes and Baer (1977) stressed the importance of generalization of treatment effects by stating that "a therapeutic behavioral change, to be effective, often…must occur over time, persons, and setting, and the effects of the change sometimes should spread to a variety of related behaviors" (p. 350). Response variability is no exception. Promoting response variability in specific situations might result in increases in reinforcement and opportunities for language development for the individual, but it is likely that those gains will be limited to the specific situations that were taught. Generalization of response variability, however, could result in more wide-spread

increases in reinforcement and opportunities for language and social development.

Generalization of response variability to appropriate situations (e.g., situations in which variability is adaptive) would be desirable as a means for individuals to contact more sources of reinforcement across a variety of settings and behaviors. For example, if a child with autism is able to vary his responses in a variety of play situations, he will likely contact even more reinforcement than if he varied his responding in one play situation but continued to engage in rigid and repetitive responding across other play situations. Engaging in a variety of play behaviors may also result in contact with sources of automatic reinforcement. That is, by varying play behaviors in different settings and situations, an individual may engage in a new behavior that is automatically reinforcing, thus increasing sources of and the amount of reinforcement. Further, achieving generalization of variability of play behaviors would allow for capitalization on play for teaching language and social interaction in a variety of play situations and contexts. If an individual is able to engage in varied play with multiple items or sets, instructors can capture more learning opportunities than if the individual only varies his or her behavior with the one play set that was trained. Thus, it would be beneficial to promote the generalization of variability to a range of play situations.

Over the past several decades, applied behavior analytic researchers have presented many different strategies for promoting generalization of treatment effects. One strategy that seems promising for promoting generalization of response variability is programming common stimuli into the teaching environment (Stokes & Baer, 1977). One way to achieve this would be to bring variability of play behavior under discriminative

control. Then, the discriminative stimulus for variability could be introduced to untrained play situations to evoke variability of play behavior. By programming a stimulus that can be common across the training environment and other environments, it may be possible to promote higher levels of variability of play behavior across multiple play situations.

Another generalization strategy that might be effective is to train multiple exemplars. In the context of variability, this might entail training variability of play behaviors in multiple different play situations during the same period of time. For example, by teaching a child to vary his behavior when playing with a farm toy, a kitchen toy, and a castle toy, he will be more likely to vary his behavior when playing with other toys such as cars and dinosaurs. It may also be possible to combine methods for promoting variability, such as programming common stimuli while conducting multiple exemplar training. Training multiple exemplars can result in a learned history of reinforcement for variability across multiple contexts. Programming for common stimuli can establish a signal, so to speak, that reinforcement is available for varied behavior across multiple contexts. Combining these two strategies may be effective at evoking varied behavior in untrained contexts.

In summary, individuals with autism often engage in invariant and repetitive play when compared to their typically developing peers. Engaging in repetitive play behavior may limit an individual's access to reinforcement and opportunities for promoting social interaction and language. Therefore, it is important to promote varied play behaviors. Response variability is an operant dimension of behavior that can be directly reinforced, and multiple strategies have been demonstrated to be effective at increasing response

variability with individuals with autism. However, little is known regarding the generalization of variability to untrained responses. Below, we summarize research on procedures for increasing response variability and promoting generalization of variability as they relate to play with children with autism. We then present the purpose of our study, along with our specific research questions. Next we describe our method for this study, followed by the results. Finally, we provide our conclusions and discussion of the implications of this research.

CHAPTER II

LITERATURE REVIEW

There is a growing body of literature on response variability that includes a range of studies from the basic, translational, and applied behavior analytic literature. Response variability has been investigated across multiple species and multiple response topographies. Below we present a review of the literature on studies demonstrating that variability is an operant dimension of behavior and studies demonstrating the use of different reinforcement procedures for increasing response variability with individuals with autism and other related disabilities. We focus primarily on studies that investigated procedures for increasing variability of play behavior. Finally, we review studies that investigated procedures for promoting generalization of response variability.

Variability as an Operant Dimension of Behavior

Pryor, Haag, and O'Reilly (1969) laid the groundwork for studying reinforcement of response variability by increasing novel behaviors emitted by a porpoise through direct reinforcement. The trainers only provided reinforcement if the porpoise engaged in a new behavior that had not been seen during previous sessions. Only one form of response was reinforced per session. Initially, the trainers reinforced responses that the porpoise was known to emit from before the experiment. After a few sessions, the trainers had reinforced a majority of different behaviors within the porpoises known repertoire. At this point, the trainers observed that the porpoise began to emit truly novel behaviors that she had never emitted before. As the experiment continued, the porpoise began to emit

novel behaviors that were complex and outside what had been observed for that species of porpoise. The results of this experiment suggested that reinforcement could be used to produce novel or creative behavior.

Goetz and Baer (1973) extended the findings of Pryor et al. (1969) by investigating the effects of differential reinforcement on novel forms of block building by three typically developing preschoolers. The researchers provided reinforcement contingent on block building forms that had not previously been emitted within that session. They found that, for all three participants, the differential reinforcement of new forms procedure resulted in increases in form diversity. The researchers then reversed the contingency such that block forms were only reinforced if they were the same as previous forms and they saw a decrease in the number of new block forms built each session. When the researchers returned to the differential reinforcement of new behaviors, they saw an increase in the number of new forms per session. These results suggest that not only can reinforcement result in increases in novel behavior, but that the emission of novel or repeated behavior is sensitive to changes in reinforcement.

Page and Neuringer (1985) confirmed that response variability is an operant dimension of behavior by demonstrating that variability was sensitive to contingencies of reinforcement and could be brought under discriminative control. Through a series of six experiments, Page and Neuringer demonstrated that variability of sequences of left and right key pecks emitted by pigeons was increased through direct contact with reinforcement contingencies for variability. They found that the pigeons would meet the variability contingency even when a sequence needed to differ from the previous 50

sequences in order to result in reinforcement. They also found that permitting variability without requiring it did not result in response variability; variability was only observed in conditions where reinforcement was contingent on variability. In the final experiment, Page and Neuringer demonstrated that response variability could be brought under discriminative control.

Although not an exhaustive list, the results of these studies serve to represent the literature base that demonstrates that response variability can be directly influenced by environmental events (i.e., antecedents and consequences). Specifically, response variability can be increased through contingencies of reinforcement and can be brought under discriminative control. Since Page and Neuringer's (1985) landmark study, much work has been done to further investigate the effects of antecedent and consequent events on response variability. One area of study that has been growing over the past decade is the application of reinforcement procedures to increasing response variability emitted by individuals with autism and other disabilities.

Increasing Response Variability with Individuals with Autism and Related Disorders

The results of studies like Goetz and Baer (1973) and Page and Neuringer (1985) have implications for the treatment of individuals with autism. One of the defining characteristics of autism is the presence of excessive repetitive behaviors (APA, 2013). In other words, many individuals with autism engage in invariant behavior to the degree that it is maladaptive. Several studies have demonstrated the effects of different reinforcement

procedures on increasing response variability, and thus decreasing invariant behavior, emitted by individuals with autism. Three reinforcement procedures that have been shown to be effective at increasing variability are percentile schedules of reinforcement, differential reinforcement, and lag schedules of reinforcement.

Percentile Schedules of Reinforcement

In a percentile schedule of reinforcement, the criterion for reinforcement is frequently recalculated based on the behaviors recently emitted by the individual. That is, the most recent behaviors are ranked and ordered along some dimension and then the criterion for reinforcement for future behaviors is set somewhere along that ranking. For example, if a researcher is trying to increase duration of a lever press, the most recent 10 responses would be ordered from highest to lowest durations. If the reinforcement contingency were set at 50% of the previous durations, then the individual would need to emit a response whose duration is longer than at least 50% of the previous responses. Percentile schedules of reinforcement are commonly used in the basic literature, and have been used to increase response variability with pigeons (Machado, 1989). In a translational study, Miller and Neuringer (2000) evaluated the effects of percentile schedules of reinforcement on increasing responses variability emitted by individuals with autism.

In their study, Miller and Neuringer (2000) evaluated the effects of percentile schedules of reinforcement on increasing response variability of sequences of left and right button presses with individuals with autism in the context of a computer game. This study included five adolescents with autism, as well as two control groups containing five

college-aged adults and four young children who had no physical or intellectual disabilities. The effects of the percentile schedules of reinforcement were evaluated using an ABA design and all three groups of participants experienced the same three conditions, where sets of four presses across two buttons resulted in reinforcement. During both A conditions, reinforcement was delivered independent of varied responding. That is, reinforcement was delivered on a probabilistic schedule such that each set of responses had a 50% chance of resulting in access to the putative reinforcer. During the B condition, the percentile schedule of reinforcement was in place. After each set of responses, the computer would compare the sequence of left and right key presses to the previous 20 sequences and would then generate a relative frequency value. This value corresponded to how often the participants emitted a given sequence of key presses. Sequences that had a higher value were those that were emitted more frequently. The reinforcement contingency was set at the 11^{th} relative frequency value meaning that only sequences whose relative frequency value was less than the 11^{th} relative frequency value would result in reinforcement. This resulted in a contingency that only reinforced infrequent responses. The schedule of reinforcement changed following every set of responses, thus promoting variability of sequences of button presses in order to access reinforcement. Miller and Neuringer found that the percentile reinforcement schedules resulted in higher levels of variability as compared to the probabilistic conditions across all three groups of participants.

Although the researchers found that percentile schedules of reinforcement were effective at increasing response variability, this procedure may not be practical for

increasing variability in the applied setting. Implementing percentile schedules of reinforcement typically requires the use of computers to constantly evaluate levels of variability and update the reinforcement contingency from trial to trial. As such, this would be a cumbersome and difficult procedure to use to increase response variability in a clinical setting. Galbicka (1994) described a method for using pencil and paper to calculate percentile schedules in the applied setting. However, this procedure was designed for directional shaping of a specific response, rather than shaping variability within a response class. Although this may someday be a viable option for reinforcing varied behavior in the applied setting, it has not yet been refined for use in this context.

Differential Reinforcement

Another reinforcement procedure that has been used to increase response variability is differential reinforcement of novel or different behavior. Differential reinforcement of novel or different behavior entails providing reinforcement for the first time a response is emitted and then placing all responses that have been previously emitted on extinction. Differential reinforcement can be applied within or across sessions. When differential reinforcement of novel or different behavior is used within sessions, individual response topographies can result in reinforcement during each session throughout a study. When differential reinforcement is used across sessions, individual response topographies are only reinforced once within an entire experiment. Differential reinforcement has been used to increase variability of mand frames (Betz, Higbee, Kelley, Sellers, & Pollard, 2011; Sellers, Kelley, Higbee, & Wolfe, 2012), martial arts techniques (Harding, Wacker, Berg, Rick, & Lee, 2004), selection of activities

(Cammilleri & Hanley, 2005), and play behaviors (Goetz & Baer, 1973; Holman, Goetz, & Baer, 1977; Lalli, Zanolli, & Whon, 1994; Newman, Reinecke, & Meinberg, 2000).

Holman et al. (1977) conducted a follow-up study to Goetz and Baer (1973) where they investigated the effects of differential reinforcement of novel behavior on increases in novel responding across two experiments. Holman et al. also investigated the extent to which increases in novel behavior would generalize to new tasks. In Experiment 1, the researchers investigated the effects of differential reinforcement on increasing novel painted forms (e.g., painting at an easel) using an ABAB design with two typically developing preschoolers. The experimenters also tested for generalization to a response form that was not exposed to the differential reinforcement procedure (i.e., block building). During baseline, responses did not result in programmed reinforcement. During differential reinforcement, the researchers provided enthusiastic and descriptive praise for new painted forms within each session (e.g., "That was different!"). All block building sessions occurred concurrently with the painting sessions, but were under baseline conditions. The researchers replicated the effects of the differential reinforcement procedure from Goetz and Baer, but found that the increases in form diversity and emission of novel forms did not generalize to block building. The researchers hypothesized that this may have been due to the dissimilarity between the two different types of responses (i.e. painting and block building).

Therefore, the researchers conducted Experiment 2 in which they applied the differential reinforcement procedure to one response form (i.e., drawing) and then assessed generalization to a similar response form (i.e., painting) and two dissimilar

response forms (i.e., building with blocks and Legos). This experiment included three typically developing preschoolers. A multiple-baseline-across-participants design was used and the procedures for baseline and differential reinforcement were the same as described above. The differential reinforcement procedure was applied to forms drawn using a felt-tip pen, and the other three responses, painting, block building, and building with Legos, remained in baseline conditions throughout the study. The researchers found that the differential reinforcement procedure again resulted in increases in form diversity and novel forms drawn using the felt-tip pen. They did observe generalization of form diversity to the similar task, painting, but did not observe generalization of form diversity to the dissimilar tasks, block and Lego building. The researchers concluded that differential reinforcement of novel behavior can result in increases in form diversity and emission of novel forms across sessions, but that generalization of these effects may be limited to topographically similar response forms.

Differential reinforcement has also been used to target response variability of other forms of play, including thematic toy play. Lalli et al. (1994) used a differential reinforcement procedure to increase the response variability of toy play with two individuals with developmental delays. The researchers used a multiple-baseline-across-participants design and implemented the design twice such that the procedures were implemented with two different toys for each participant. Variability was measured as the cumulative number of novel toy play actions emitted across the entire study. During baseline, the researchers provided the participants with a single toy (e.g., an airplane) and said, “Let’s play.” The researchers provided praise for appropriate sitting, but did not

provide any contingent consequences for toy play. The researchers then conducted a single training session where they taught the participants to engage in one response with the toy (e.g., taught the participants to move the airplane in a horizontal motion). Following the training session, the researchers implemented the differential reinforcement condition where they provided descriptive praise for the first three occurrences of a play action. If that play action was repeated a fourth time within the session, it did not result in reinforcement (was placed on extinction). During baseline, neither participant engaged in any toy play. Following training and introduction of the differential reinforcement procedure, the cumulative number of untrained toy play actions increased for both participants to between six and nine different actions. The result of this study further demonstrated the use of differential reinforcement as a method for increasing response variability.

These studies demonstrate that differential reinforcement of novel or different behavior can result in increases in response variability of play behaviors. By definition, these differential reinforcement procedures involve providing reinforcement for each different or novel response a limited number of times. That is, once a response topography has been emitted that is different from all previous topographies, it will no longer result in reinforcement if emitted again within that session or experiment. Although this can result in increases in response variability, it may limit access to reinforcement when an individual is varying his or her responses across multiple different responses but isn't necessarily emitting new or novel responses within a session or across sessions.

According to Neuringer (2004), the term "variability" suggests a continuum from repetitive to random responding, both of which include at least some repetition of responses. By reinforcing only when a new or novel behavior is emitted within a session, researchers are programming a type of variability that does not allow for individuals to contact reinforcement for any level of repetition. It may be acceptable and even desired to program for reinforcement of variability that involves some amount of repetition of responses where individuals could contact reinforcement for responses even if they have been emitted within that session. This would allow for the reinforcement, and thus strengthening, of multiple individual responses as well as response variability. Additionally, it may more resemble the natural environment to allow for reinforcement of responses that have already occurred within a session. For example, when a child is playing with a toy farm set, it would be acceptable for him to walk the cow, feed the horse, make the rooster crow, and then walk the cow again. In this example, if we were only providing reinforcement for new behaviors, the child would only have accessed the reinforcer three times. But, if we were programming for reinforcement following varied behaviors that have already occurred within a session, he would have accessed the reinforcer four times.

Lag Schedules

Lag schedules of reinforcement may be better suited than differential reinforcement for promoting variability across a set or class of responses. Lag schedules of reinforcement specify a reinforcement contingency where a response is only reinforced if it is different from some specified number of previous responses. A lag *x* schedule of

reinforcement stipulates that a response can only result in reinforcement if it differs from or varies from *x* number of previous responses. For example, in a Lag 3 schedule of reinforcement, a response can only result in reinforcement if it varied from the previous three responses. This means that the individual would need to vary his or her responding across at least four different response forms in order to access reinforcement after each response. Lag schedules allow an individual to continue to access reinforcement for response topographies they have already emitted within a session while also directly promoting and reinforcing varied responding. Lower lag values will allow for more reinforcement of repeated responses than higher lag values. Thus, lag schedules can support a range of variability. Lag schedules have been demonstrated to be effective at increasing response variability across a wide range of responses, including sequences of button pressing (Murray & Healy, 2013), intraverbal responding (Contreras & Betz, 2016; Lee, McComas, & Jawor, 2002; Lee & Sturmey, 2006, 2014; Susa & Schlinger, 2012; Wiskow & Donaldson, 2016), tacting (Heldt & Schlinger, 2012); phonemes (Esch et al., 2009; Koehler-Platten, Grow, Schulze, & Bertone, 2013), manding (Brodhead et al., 2016), and play behavior (Baruni et al., 2014; Harris, 2016; Napolitano, Smith, Zarcone, Goodkin, & McAdam, 2010).

Baruni et al. (2014) used lag schedules to increase variability of toy play with three elementary aged children with autism and other intellectual disabilities using a multiple baseline design. The type of play under investigation was the manipulation of a single toy item (e.g., a toy car). Baruni et al. measured variability as the cumulative number of novel toy play responses across sessions, where they scored a response as

novel if it had not previously been emitted throughout the entire study. During baseline, the researchers gave the participants their assigned toy and said, “play.” There were no programmed consequences for toy play responses. During the lag conditions, toy play responses resulted in access to a preferred edible item according to the lag contingency. All three participants underwent baseline and Lag 1 conditions, and two participants were exposed to a Lag 2 condition following the Lag 1 condition. The researchers found that all three participants emitted some novel responses during baseline (between six and nine responses) but that there was a marked increase in the emission of novel responses during the Lag 1 condition for two participants (between 11 and 17 novel responses) and a modest increase for the third participant (six novel responses). The researchers saw limited increases in novel responding for the two participants exposed to the Lag 2 condition (one novel response for each participant). The results of this study suggest that a Lag 1 schedule of reinforcement was effective at increasing novel play actions with specific toys. The limited number of novel responses emitted during the Lag 2 condition suggests that increasing the lag schedule may not result in further increases in novel play. The researchers suggested that this may have been due to a ceiling effect. That is, there may have been a limited number of possible actions to do with the toys.

The results of the Baruni et al. (2014) study provide further demonstration that lag schedules can be used to increase novel responding across time, but they do not speak to the ability of lag schedules to promote variability around existing responses. Under a Lag 1 schedule of reinforcement, it is likely that the participants emitted and received reinforcement for responses that had previously been emitted. Displaying novel responses

emitted across sessions does not provide information on variability across previously emitted responses within each session. Having an additional measure such as number of different responses emitted each session or number of responses that met the lag contingency would provide more complete information on the extent to which lag schedules increase variability of toy play. Another limitation to note is the type of play responses that were reinforced throughout the course of this study. In this study, a response was reinforced as long as it was different from the previous response, regardless of how appropriate or contextual it was. For example, if the participant had put the toy in his mouth, this was reinforced if it was different from the previous action. Thus, it is possible that the procedures used in this study resulted in increases in inappropriate toy play, along with increases in varied appropriate toy play.

In an unpublished dissertation, Harris (2016) extended the Baruni et al. (2014) study and investigated the effects of a Lag 1 schedule of reinforcement plus prompting on variability of toy play in preschool children with autism using a multiple-baseline-across-participants design. The form of play under investigation was play with toy sets that involved a large 3D backdrop with multiple corresponding toy pieces (e.g., Little People™ farm set). Thus, Harris extended the Baruni study by investigating more complex forms of play. Also, Harris looked specifically at increasing variability of appropriate play actions. The primary variability measures were the number of different appropriate play actions per session and the number of appropriate play actions that met the lag criterion per session. A play action was considered different if it differed along one dimension (i.e., character, action, and location) from all play actions emitted within

that session. In order to meet the Lag 1 criterion, a play action needed to differ from the previous play action along two dimensions. During baseline, the researchers provided no prompts or reinforcement for play actions. The intervention consisted of a Lag 1 schedule of reinforcement plus a prompting procedure to evoke play actions. If a play action was different from the previous action along two dimensions, regardless if it was prompted or not, the researcher provided access to a preferred edible or sound bite of a preferred song. The researcher allowed each participant 30 seconds at the beginning of each session to engage in play actions. If the participant engaged in play actions that met the lag schedule, the researcher provided reinforcement and extended the 30 second interval. If the participant did not engage in a play action that met the lag schedule, the researcher prompted a series of three play actions that each met the lag schedule and thus resulted in access to reinforcement. Throughout both conditions, the researcher conducted generalization probes to two different play sets using baseline procedures. Following the Lag 1 plus prompting condition, the researcher conducted maintenance probes at 1 day and 2 weeks post intervention. Play actions were reinforced on an FR1 schedule during maintenance sessions.

Harris (2016) found that during baseline, none of the participants emitted different appropriate play actions or appropriate actions that would have met a Lag 1 schedule of reinforcement. When the Lag 1 plus prompting procedure was put in place, the number of different appropriate play actions and number of appropriate play actions that met the lag requirement increased for two participants and increased for the third once modifications were made to his reinforcement procedures (i.e., switched from tokens to delivery of

music contingent on meeting the lag schedule). Following the lag plus prompting intervention, Harris saw that the number of different appropriate play actions maintained in the absence of the lag schedule and prompting at both the 1 day and 2 week probes. Harris also found that the number of appropriate play actions that met the lag criterion maintained for all three participants at the one-day probe, and maintained for two participants at the 2-week probe. Harris saw limited generalization to the other two play sets throughout the study. One of the participants emitted different play actions only during the final generalization probe for each of the sets. The other two participants emitted some different play actions during generalization probes throughout the study, but levels were lower than those seen in the intervention set.

The results of the Harris (2016) study are important, as this was the first study to target increasing variability of *appropriate* and complex toy play. As discussed above, individuals with autism sometimes engage in invariant behavior to the point that it is maladaptive. Procedures such as lag schedules of reinforcement are effective at reducing invariant behavior. However, in the context of play, it may not be sufficient to simply increase varied play behavior infinitum. In situations where response variability is adaptive, levels of variability that are desirable might be different depending on the type of response. For some types of responses, such as manding or intraverbal responding, it might be desirable for a person to vary his or her responding across a relatively small number of responses. The limits of which behaviors are appropriate for the person to vary across are restricted by the response class itself. For other types of responses, such as play, it may be more desirable for a person to vary his or her responding across a large

number of responses. However, there are still limits to which behaviors are adaptive in play. For example, increases in variability of inappropriate play behavior (such as putting a toy car in your mouth) may still be maladaptive. Thus, it is important to target variability of play behavior within certain parameters, such that we are increasing variability of appropriate play behaviors.

Although there is a growing body of literature that supports methods for increasing response variability, there are surprisingly few studies that have investigated generalization of the effects of these interventions. Some studies assessed for generalization of variability to new locations and new people and observed that variability did generalize (e.g., Betz et al., 2011, Brodhead et al, 2016; Harding et al, 2004; Sellers et al., 2016). In many studies, the response or response class under investigation was the same as in treatment. Some studies also evaluated whether variability generalized to different responses, but generalization was rarely seen and was limited when it was observed (e.g., Harris, 2016; Holman et al., 1977; Napolitano et al., 2000). This is not entirely surprising, however, because these studies did not include procedures for promoting generalization of variability. As outlined by Stokes and Baer (1977), assessing for generalization without explicitly programming for it can provide important information on the extent to which the results of certain procedures may or may not generalize. When results are not observed to generalize, Stokes and Baer stress the importance of programming for generalization and provide descriptions for multiple methods to do so. To this end, studies that assessed generalization provide important information regarding the generalization of variability. That is, they suggest that

variability is unlikely to generalize to new responses without programming for it to do so. Thus the next step is to implement procedures to specifically program for the generalization of variability to new responses.

Generalization of Variability

Few studies have directly programmed for generalization of variability across responses (Brodhead et al., 2016; Miller, 2014; Parsonson & Baer, 1978). Two methods for promoting generalization of variability that seem promising are multiple exemplar training and programming common stimuli. Multiple exemplar training involves training a response across multiple examples of the stimuli associated with the response. In the context of variability of play behavior, multiple exemplar training might entail training an individual to engage in varied behavior with multiple sets of play materials. This would result in a learned history of engaging in varied behavior across multiple different response classes, and thus may promote the individual engaging in varied behavior with untrained play sets. To date, one study has investigated the effects on multiple exemplar training on promoting response variability.

Parsonson and Baer (1978) evaluated the effects of multiple exemplar training on generalization of novel tool use during problem solving tasks. The researchers implemented differential reinforcement of novel tool use across multiple examples of tools in problem solving tasks and then tested for novel tool selection in untrained tasks. This study included five children, three typically developing preschoolers and two elementary-age children with “language and behavioral problems.” Improvisation

training took place in the context of three different problem solving tasks; pounding on a pop-up peg toy, sorting marbles, and lacing a shoe. For each task, items were missing that were integral to completing the task, thus creating the opportunity for problem solving. The items required to complete these tasks that were missing were a hammer, storage containers, and a shoe lace respectively. Therefore, each task required the participant to find a different item or combination of items to serve the same function as the missing item.

There were two types of sessions conducted in this study, training sessions and probe sessions. One group of items for each task was present during teaching sessions and a different set of items was available during probe sessions. For each set, the researchers provided five to seven items that could potentially be used as replacements for the missing items (e.g., a small shoe for the pop-up peg toy) along with 10 distractor items that would not be effective as replacements (e.g., a string for the pop-up peg toy). The primary measure in this study was the cumulative number of new improvisations across sessions, where the use of an item was considered a new improvisation if it was effective at accomplishing the task and had not previously been used throughout the entire study. The researchers only presented the data for the probe sessions. Thus, the researchers only presented data on generalization of the effects of the training to the probe sets. Probe sessions occurred throughout the entire study and always followed training sessions. During probe sessions, use of new items to complete the tasks resulted in the researchers providing mild praise in a “matter-of-fact” tone. During baseline training sessions, the participants were presented with the task along with one common

item to use to complete the task (e.g., a wooden hammer for the peg toy). If the participant completed the task, the researcher provided enthusiastic praise. During improvisation training sessions, the researchers presented new items in each session with which the participant could complete the task and then provided enthusiastic praise when the participant successfully used the item. New items were introduced one at a time across successive training sessions.

The researchers evaluated the effects of the multiple exemplar training using a multiple baseline across behaviors design (i.e., across tasks) for three participants and a combined multiple baseline across behaviors and participants design for two participants. In general, the researchers saw that training multiple examples of novel tool use in one task resulted in generalization of novel tool use in the probe sessions for that task, but did not result in generalization to the untrained tasks. For one participant, the researchers implemented multiple exemplar training in the peg toy task only. They saw that the training resulted in increased novel tool selection during the probe sessions for the peg toy task but not for the shoelace or marble task. For two participants, the researchers implemented multiple exemplar training in two tasks. They saw increases in novel tool use for the first task that was trained for both participants (up to six more than baseline), but levels of novel tool use emitted during training of the second task were similar to baseline (one to two responses). Novel tool use remained low in the third (untrained) task. For the remaining two participants, multiple exemplar training was implemented across all three tasks in a staggered fashion across both tasks and participants. The researchers saw increases in novel tool use following multiple exemplar training in all

three tasks for both participants.

The results for the five participants, taken together, suggest that the multiple exemplar training resulted in generalization of new improvisations within the tasks that had been trained but did not result in generalization to tasks that had not been trained. In other words, the multiple exemplar training resulted in generalization within response classes, but not across response classes. The results of this study suggest that multiple exemplar training may be a promising method for promoting generalization of varied behavior. However, this is a tentative and preliminary conclusion. Parsonson and Baer (1978) only saw generalization within responses classes and not across. However, they only implemented multiple exemplar training within responses classes, and not across. It is possible that conducting multiple exemplar training for variability across response classes may result in generalization of variability across response classes. That is, simultaneously training variability in multiple response classes may result in increases in other responses without explicit training. In summary, the results of Parsonson and Baer suggest that multiple exemplar training may be promising method for promoting generalization of variability.

Another possible method for promoting generalization of variability of play behavior is to program common stimuli across the play materials that are trained and those that are untrained. This could be achieved by bringing variability of play under discriminative control through discrimination training. The discriminative stimulus that signals the availability of reinforcement for varied behavior could then be presented in the presence of untrained play materials to evoke variability. Thus, the discriminative

stimulus for variability would be a common stimulus across trained and untrained play materials and thus may promote generalization of response variability to the untrained play materials. Two studies to date have evaluated the effects of discrimination training on promoting generalization of response variability.

In an unpublished dissertation, Miller (2014) investigated the effects of using discrimination training to establish stimulus control over variability in a play task and then tested whether the discriminative stimulus for variability would evoke variability in another play task. This study included three elementary aged children with autism. Response variability was trained in the context of block building with two participants (Reed and Pete) and placing pegs in a peg board for the third participant (Don). Generalization was assessed in the context of a painting activity for all three participants. Variability of block structures, peg patterns, and painting patterns were measured in two ways. *U*-value was calculated for the first phases of the study. *U*-value determines the distribution of responses across all possible responses. In order to calculate *U*-value, there needs to be a known finite number of possible responses from which to vary (e.g., 17 possible block or peg structures and 16 possible painting patterns). This allows for the comparison of the frequency of one response to the frequency of all other possible responses. Higher *U*-values indicate higher levels of variability and lower *U*-values indicate lower levels of variability. The second measure of variability was the percentage of trials meeting a Lag 3 contingency for each phase. For each trial, the response was compared to the previous three responses. If a response form differed from the previous three, it was considered to meet the Lag 3 contingency and if a form was the same as any

of the previous three, it was not considered to meet the Lag 3 contingency.

Miller (2014) implemented a reversal design across multiple phases. During the continuous reinforcement (CRF) phase, the therapist presented the discriminative stimulus (S^D), "Build something" and complete responses resulted in reinforcement. During the vary phase (VAR), the therapist presented the S^D, "Build something different" and responses were reinforced according to a Lag 3 schedule of reinforcement. During the repetition phase (REP), the therapist presented the S^D, "Build the same" and responses resulted in reinforcement according to a Rep 3 schedule (i.e., if they were the same as one of the previous three response). For all VAR and REP sessions, an error correction procedure was in place where, following three trials where the participant did not meet the contingency, the researcher modeled a correct response and had the participant imitate. During the discrimination training phase (ALT), the therapist alternated between VAR and REP trials in a multiple schedule. Each trial had an equal probability of operating under the VAR or REP procedures (including the associated S^Ds and schedules of reinforcement). During the S^D absent phase (SDA), the therapist alternated between VAR and REP trials in a mixed schedule where the S^Ds were not presented. During the generalization phase (GEN), the procedures from the ALT phase were in place but the participant was asked to engage in the painting task instead of the block building or peg task.

The U-value statistic was calculated for the first five phases (CRF, VAR, REP, VAR, REP) for Don and Reed and the first seven phases for Pete (CRF, VAR, REP, VAR, REP, VAR, REP). Miller saw that, for all participants, the *U*-value was low during

the CRF phase for all participants, was higher during the VAR phases and lower during the REP phases (relative to the VAR phases). The results for this measure demonstrate that the Lag 3 schedule of reinforcement resulted in increased variability and the Rep 3 schedule resulted in reduced variability for all three participants. However, Pete required additional exposure to the VAR and REP phases before his responding came under control of the different contingencies. The percentage of trials meeting the Lag 3 contingency was presented for all phases for all participants. The data for the first five (Don and Reed) and seven (Pete) phases reflect the patterns seen in the *U*-values for these phases. That is, for each participant, more trials met the Lag 3 contingency during the VAR phases than during the CRF or REP phases. During the discrimination training phases, Miller saw that response variability came under discriminative control for all three participants. In general, the percentage of trials that met the Lag 3 contingency were higher under the VAR trails and lower under the REP trials during the ALT phases. The percentage of trials meeting the Lag 3 contingency in the VAR and REP trials during the SDA phases (where the S^Ds were not present) converged and were roughly the same. During the GEN phase the percentage of trials meeting the Lag 3 contingency were either lower in the VAR trials that in the REP trials (Don) or the difference between VAR and REP trials was minimal (Reed and Pete). Thus, generalization of the discrimination between VAR and REP trials to a new task was not observed for any of the participants.

There are several limitations to this study that might have contributed to the failure of responding to generalize to a new task. First, levels of variability during the VAR conditions, although elevated compared to the REP conditions, were still low. The

contingencies of reinforcement during the VAR condition may have been set too high. In order to contact high levels of reinforcement under a Lag 3 contingency, the participants needed to emit four or more different responses. Miller reported that during the CRF phase, two participants emitted three different responses across all CRF sessions and one participant emitted just two different responses. This may have accounted for the generally low levels of variability observed during the VAR phases. The Rep 3 contingency may have been too high as well. In a Rep 3 contingency, a response will result in reinforcement if it is the same as one of the previous three responses. This contingency allows for some amount of response variability to occur and still contact reinforcement. Therefore, the Rep 3 contingency may have resulted in the incidental reinforcement of variability. A lower rep contingency may have resulted in even greater differentiation between the VAR and REP conditions.

A second limitation to this study was the results and nature of the discrimination training. Even though the patterns of responding across the ALT and SDA phases suggest that variability was brought under discriminative control for all participants, these did not appear to be strong discriminations. For example, the most drastic discrimination for Don was observed during the first ALT phase where about 5% of the trials met the Lag 3 contingency during the REP condition and about 50% of the trials met the Lag 3 contingency during the VAR condition. One reason that strong discriminations were not observed might have been the nature of the S^D used to signal the availability of reinforcement for varied or repetitive behavior. The researcher used vocal S^Ds (e.g., "build something different"). One drawback to using a vocal S^D is that it is only present

for a very brief amount of time; once the instruction is given the S^D is gone. It is common in discrimination training to use a more salient S^D, such as a colored card, that can remain present throughout the entire condition. Using a more salient S^D may have resulted in stronger discriminations with these participants.

Another reason that the discrimination training did not result in strong discriminations may have been the procedures of the discrimination training itself. During the ALT phases, a multiple schedule was in place that involved alternations between VAR and REP trials. That is, each individual trial had a 50% chance of being reinforced on a Lag 3 or Rep 3 contingency. It is possible that alternating the VAR and REP contingencies on a trial-by-trial basis may have been too rapid for the participants to contact the contingencies in the presence of the associated S^Ds. It may have been more effective for the VAR and REP conditions within the multiple schedule to consist of multiple trials each. This may have resulted in the participant contacting reinforcement associated with the relevant S^D multiple trials in a row and thus may have resulted in a stronger discrimination.

A recent study published by Brodhead et al. (2016) addressed some of these limitations to bring variability of mand frames under discriminative control with preschoolers with autism. The researchers also investigated the effects of script training combined with lag schedules of reinforcement on increasing response variability. As the response under investigation was mand frames (e.g., “I want –” or “Can I have –”), all sessions took place during a brief snack time. The primary measure was the number of different mand frames emitted each session, where a frame was considered different if it

had not been emitted within that session. During baseline, all mand frames were reinforced with a small piece of the requested snack. Two participants primarily emitted just one mand frame across most baseline sessions, and the third participant emitted up to three different mand frames, but emitted one or two mand frames per session for most baseline sessions.

Following baseline, discrimination training was implemented where the researchers alternated across sessions where the lag contingency was in place or a repetition contingency was in place. All discrimination training sessions included script training, where the researchers prompted the participants to point to and read a written script of the different mand frames, which were then reinforced according to the appropriate contingency. During sessions under the lag conditions, a green placemat was present to signal the availability of reinforcement for varied responding. The lag schedule used for each participant was based on the number of different frames emitted during baseline and was set at a Lag 2 for two participants and a Lag 3 for the third. During sessions under the repetition contingency, a red placemat was present to signal the availability of reinforcement for repeated mand frames. During these sessions, only the frame “I want _____” resulted in reinforcement. The researchers found that the number of different mand frames emitted each session increased in the lag schedule component and remained low around one or two mand frames per session in the repetition component. The scripts were systematically faded out until they were no longer present in the sessions, meaning that only the red or green placemats were present during sessions. The lag and repetition contingencies were still in place and the separation between

responding in the lag component and the repetition component maintained in the absence of the scripts, suggesting that variability of mand frames had been brought under discriminative control.

Next, the researchers removed the reinforcement contingencies but kept the colored placemats in place. That is, all mand frames were reinforced on an FR1 contingency in the presence of both the red and green placemat, regardless of variability or repetition. They found the separation between varied responding in the presence of the red and green placemats remained for two participants, suggesting that the placemats were functioning as discriminative stimuli evoking either varied or repetitive responding. The researchers then tested for generalization of the effects of the placemats to a different setting, and saw high levels of variability in the presence of the green placemat and low in the presence of the red. For the third participant, when the reinforcement contingencies were removed, the number of different mand frames decreased to one in the presence of both placemats. Therefore, the lag and repetition contingencies were put back in place for this participant and his responding immediately separated again to match the contingencies. The researchers then conducted generalization sessions to a new setting with the contingencies in place, and saw that variable responding and repetitive responding occurred in the respective sessions.

The results of this study have several implications. First, these results demonstrate that variability of mand frames can be brought under discriminative control. Brodhead et al. (2016) saw a strong discrimination for all three participants with high variability in the vary component and low variability in the repeat component. It is worth noting that

Brodhead et al. implemented procedures that addressed some of the limitations from the Miller (2014) study. First, they based the lag schedules off the participants' baseline levels of responding such that the participants were likely to contact the lag requirement in the vary condition. They also used a low repetition contingency (i.e., a Rep1) which resulted in low levels of variability in the repetition condition. They also used salient visual S^Ds (i.e., the green and red placemats) that were present during the entire sessions during the discrimination training. They also rotated between variability and repetition conditions on a session-by-session basis, thus ensuring that the participants contacted reinforcement multiple times in the presence of the S^Ds during the discrimination training. These modifications may have contributed to the strong discrimination that was observed between variability and repetition with these participants.

A second implication of this study is that these results suggest that the discriminative stimuli that are associated with variability might be used to promote the use of varied mand frames when the reinforcement contingencies are not necessarily in favor of variability. In other words, the discriminative stimuli may promote maintenance of response variability. This lends support to the idea that discriminative stimuli may be used to promote generalization, as maintenance of responding is a form of generalization across time. It is possible that achieving strong discrimination of response variability may also promote generalization to other responses.

In summary, two methods for promoting generalization, multiple exemplar training and programming common stimuli, have been investigated as means for promoting generalization of response variability. However, there were limitations to the

generalization observed in each study. Parsonson and Baer (1978) found that multiple exemplar training may be effective at promoting generalization of novel behavior within responses classes, but not across response classes. Miller (2014) found discrimination training was effective at bringing variability of play behavior under discriminative control, but did not observe generalization of response variability when the discriminative stimuli were applied to untrained play sets. Brodhead et al. (2016) demonstrated that variability of mand frames could also be brought under discriminative control and saw that the discriminative stimuli supported maintenance of varied responding. Each of these studies investigated one method for promoting generalization. It is possible that combining methods may result in more robust generalization of response variability. That is, combining multiple exemplar training with programming common stimuli (via discrimination training) may result in generalization of response variability to untrained situations. Therefore, the purpose of this dissertation was to investigate the effects of combining multiple exemplar training with discrimination training on promoting generalization of variability of play actions emitted by preschoolers with autism. The specific research questions for this study were:

1. Will lag schedules of reinforcement that are implemented across multiple play sets (i.e., multiple exemplar training) result in increases in the varied play behavior of children with autism, as measured by the number of responses meeting a Lag 1 or Lag 2 schedule and the number of different responses emitted each session?

2. Will discrimination training, including variability and repetition components, implemented across multiple play sets (i.e., multiple exemplar training), result in discrimination of varied play behaviors emitted by children with autism?

3. If variability of play behavior emitted by children with autism is brought under discriminative control, will the effects of the discrimination training generalize to untrained play sets?

4. If variability of play behavior is not brought under discriminative control, will multiple exemplar training alone be effective at promoting generalization of varied play behavior to untrained play sets?

CHAPTER III

METHODS

Participants

Three preschool-aged children with autism participated in this study. All three were recruited through, and were current students at, the ASSERT autism preschool. We initially recruited six participants, but screened out three during the inclusion assessment process. The inclusion assessment consisted of three 5-minute play sessions where we told the participant to "go play" with a play set that contained a large 3-D background and five figurines. The play sets used during the inclusion assessment were similar to, but were not identical to, the play sets used during the study. The therapist running the inclusion assessments did not provide any differential consequences for the participants' behavior during these sessions. We included participants in this study if they engaged with the play sets but displayed low levels of variability of appropriate play actions. A participant was considered to have engaged with the play sets if he touched any part of the play set at any point during the 5-minute play sessions. If the participant's appropriate play behavior met a Lag 2 schedule more than 10 times (using the response measures outlined below), he was not included. Also, we only included participants for whom edibles were likely to function as reinforcement for their behavior. This was determined through caregiver and clinician report. All participants were reported to respond to edibles as reinforcers.

Following the inclusion assessments, we included three participants in the study.

All three were males who attended the ASSERT autism preschool where they received early intensive behavioral intervention for 20 hours per week. Leonard was 4 years old at the time of the study and had been attending ASSERT for 12 months. Leonard also attended a public special education preschool approximately 12 hours a week. At the time of the study, Leonard was able to imitate simple motor actions and was beginning to learn to receptively identify common objects and to respond to simple one-step instructions. He requested using picture exchange, and discriminated across multiple pictures at once. Leonard had a strong echoic repertoire (for one syllable words) and was beginning to emit one word requests vocally. Leonard also responded intraverbally to certain play-related phrases (e.g., "Ready…set…. [go]"). Leonard did not engage in independent play, but would play appropriately when prompted by adults. During independent play, he engaged in repetitive movements such as picking up one toy at a time and spinning it. During the inclusion assessments, Leonard interacted with the plat sets by touching and moving the figurines around. Leonard completed one appropriate play action during one of the inclusion assessment sessions (with the house play set). Thus, his play behavior did not meet a Lag 1 or Lag 2 schedule during any of the inclusion assessments.

Arthur was 5 years old at the time of the study, and had been attending ASSERT for 7 months. Arthur attended a public special education preschool approximately 12 hours a week. At the beginning of the study, Arthur was able to imitate gross motor movements and was beginning to respond to one-step instructions. He was also beginning to receptively label common objects. Arthur requested using picture exchange, and discriminated across multiple pictures. Arthur did not have an echoic repertoire, but did

babble frequently throughout the day. Arthur would sometimes play independently with cause and effect toys, but often engaged in stereotypy (e.g., spinning in circles) during play. During the inclusion assessments Arthur interacted with the play sets by holding one or two figurines at a time. He also frequently put the figurines in his mouth. Arthur completed one appropriate play action during one of the inclusion assessment sessions (with the house play set). Thus, his behavior also did not meet a Lag 1 or Lag 2 schedule during the inclusion assessment.

Jude was also 5 years old at the time of the study, and had been attending ASSERT for 7 months. Jude attended a public special education preschool approximately 12 hours a week. At the time of the study Jude had a strong imitative repertoire, was able to follow two-step instructions, and could receptively identify common objects when presented in an array of three to five items. Jude communicated his wants and needs using three word phrases (e.g., "I want [item]") and was able to echo most words and short phrases. Jude was beginning to acquire vocal labeling and simple intraverbal responding (e.g., answering yes/no questions). When presented with play materials, Jude engaged in stereotyped play behaviors such as stacking or lining up toys and holding toys up to his eyes one by one. During the inclusion assessments Jude held the figurines and moved them around, and also touched the 3D backdrop portion of the play set. Jude's play behavior met the Lag 1 and Lag 2 schedule one time during one of the inclusion assessment sessions (with the castle).

Setting and Materials

We conducted all sessions in a small research room that was approximately 1.5 m by 1.5 m with a two-way mirror on one of the walls. The room contained a small table, two chairs, a video camera, play materials, and edibles to be used as reinforcement. We selected edibles for each participant based on caregiver report and existing preference assessment data (see Table 1). We used six play sets for each participant in this study, three of which were assigned to multiple exemplar training plus discrimination training and three assigned to the generalization condition. We used a nine-item paired stimulus preference assessment to identify each participants' preference and included the top six sets. We conducted the preference assessments based on the procedures described in Fisher et al. (1992). We placed two play sets on the small table in the research room and said, "Pick one." Once the participant touched a play set, we removed the other play set and allowed the participant 30 s to interact with the selected play set. We conducted a total of 72 trials, such that each play set was paired with each other play set twice, once on the left and once of the right. Table 2 displays descriptions of the play sets included in

Table 1

Edibles for Each Participant

Leonard	Arthur	Jude
Cheese Crackers	Gummy bears	Skittles
Doritos	Swedish Fish	Marshmallows
Oreos	Potato chips	Cheez-its
Goldfish crackers	Oreos	Goldfish
Cheetos	Cheetos	Chocolate chip (sessions 42-60) Frosting (sessions 61-137)

Table 2

Play Sets Included in the Initial Paired Stimulus Preference Assessment (Exceptions to the definition of appropriate play actions are also listed)

Play set	Figurines	Picture	Exception
Batman	Batman Robin Joker Motorcycle		NA
Castle	Knight Princess Horse Boulder		Slide under bed
Doc McStuffins	Doc Baby Lamb Helicopter		NA
Farm	Farmer Cow Pig Hay		Slide within silo
Fire station	Fireman Bot Dog Motorcycle		Hole for elevator

(*table continues*)

Play set	Figurines	Picture	Exception
House	Mom Dad Baby Car		NA
Lion King	Lion Hyena Badger Boulder		NA
Mini Mouse	Mini Daisy Cat Car		Slide
Paw Patrol	Dog Monkey Zebra Jeep		Trap door
Pirate ship	Jake Pirate Monster Boulder		Trap door Rope ladder

the preference assessment. Each play set was accompanied by four figurines and were approximately the same size and allowed for similar amounts of play actions.

Figure 1 shows the results for the preference assessments for the three participants. We conducted two separate preference assessments with Jude (both are included the Figure 1). The first assessment included the same play sets as the other two participants. During Jude's initial baseline, he was engaging in high levels of varied play actions with the fire station play set but did not engage in varied play with the other two sets. We then excluded the fire station from the pool of play sets and conducted a second preference assessment where we replaced the fire station play set with the Mini Mouse

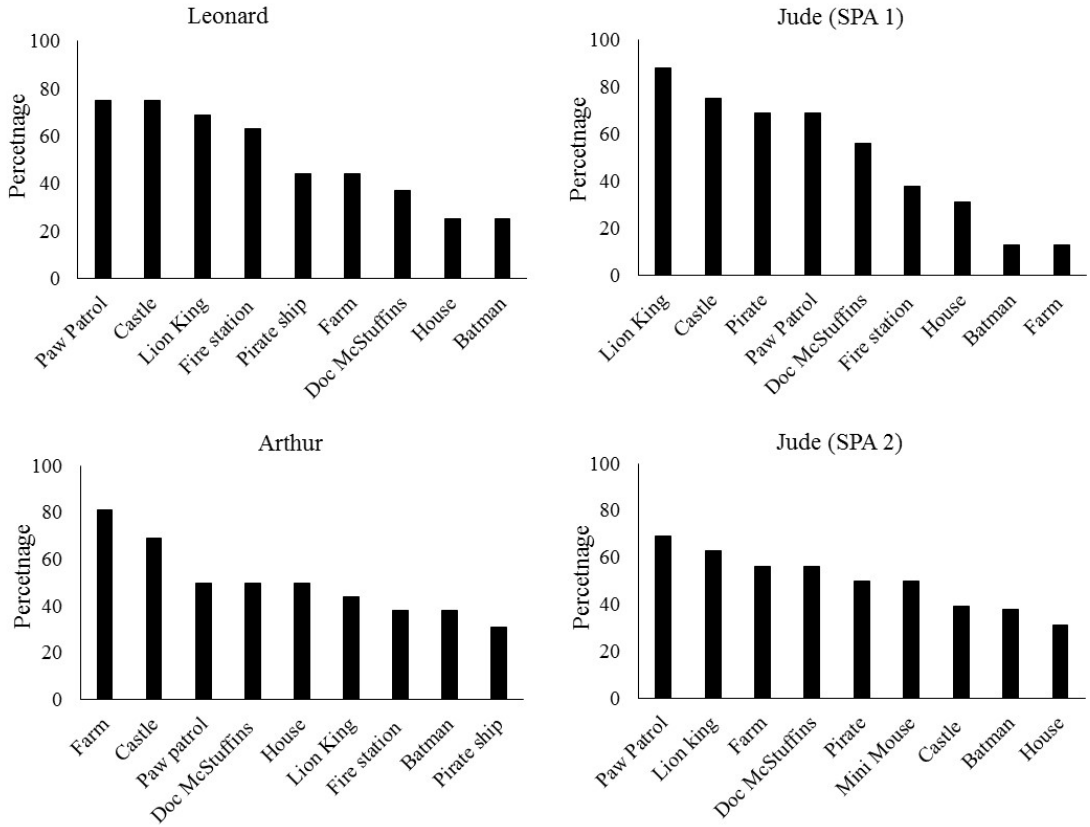

Figure 1. Percentage of trials that each participant selected each play set during the paired stimulus preference assessment.

play set. The play sets selected from this second preference assessment are the ones we used for the remainder of the study. After we identified which six play sets we would include for each participant, we assigned the play sets to the training condition or the generalization condition based on their rankings. We flipped a coin to determine whether play sets ranked with even numbers would go into the training condition (heads) or the generalization condition (tails). This allowed us to control for the effects of preference within each condition by roughly equating preference for play sets across conditions. Table 3 shows which play sets were assigned to which condition for each participant.

Response Measures and Data Collection

We had two primary measures and one secondary measure in this study. The

Table 3

Assignment of Play Sets to Conditions for Each Participant, and Symbols Used to Represent Each Play Set On the Figures

	Condition					
	Training			Generalization		
Participant	Play set	Rank	Symbol	Play set	Rank	Symbol
Leonard	Castle	2	●	Paw Patrol	1	×
	Fire Station	4	▲	Lion King	3	+
	Farm	6	■	Pirate Ship	5	◆
Arthur	Castle	2	▲	Farm	1	+
	Doc McStuffins	4	■	Paw Patrol	3	×
	Lion King	6	●	House	5	◆
Jude	Paw Patrol	1	▲	Lion King	2	◆
	Farm	3	●	Doc McStuffins	4	+
	Pirate	5	■	Mini Mouse	6	×

primary dependent variables were, (1) the number of independently performed play actions that met the lag schedule and, (2) the number of independently completed different appropriate play actions (i.e., number different) during each session. For two participants (Leonard and Arthur), a Lag 1 schedule was in place and for one participant (Jude) a Lag 2 schedule was in place. Thus, for Leonard and Arthur, the first primary dependent measure was the number of independently performed appropriate play actions that met a Lag 1 schedule, and for Jude this measure was the number meeting a Lag 2 schedule.

We defined appropriate play actions as any motor movement of a figurine (e.g., toy knight) interacting with other parts of the play set (e.g., castle with doors and a catapult, princess, horse, and boulder; Harris, 2016). Appropriate play actions involved three components; (a) selection of a figurine, (b) movement of the figurine, (c) location with respect to the play set and other figurines. Each play set included four figurines (listed in Table 2). We defined movement as the participant holding the figure and making it move. That is, the participant's hand and figure needed to be moving together in order to count as movement. For example, holding the horse and making it walk, holding the princess and making her jump, and moving the knight to launch the catapult all counted as movement. Placing or setting a figurine in a location, holding a figurine still, or dropping a figurine did not count as movement. One exception to this definition was letting go of or dropping a figurine such that the figurine falls through a hole or down a slide. Six play sets allowed for this exception (listed in Table 2). The locations within each play set were pre-specified and agreed upon by all data collectors. An

example of a play set is shown in Figure 2 with each of the locations labeled.

Given this definition, examples of appropriate play actions for a play set that includes a knight, princess, horse, and a boulder include (a) placing the knight on the horse and making the horse run, (b) moving the knight to the castle and opening the door, and (c) making the knight launch the boulder from the catapult. In each of these examples, the play action includes all three components. That is, each play action includes movement of a figurine with respect to the other components of the play set. The following would not be considered appropriate play actions because they do not include all three components; (a) holding the knight still in front of the castle, (b) rolling or bouncing the boulder on the floor away from the play set, (c) manipulating the catapult

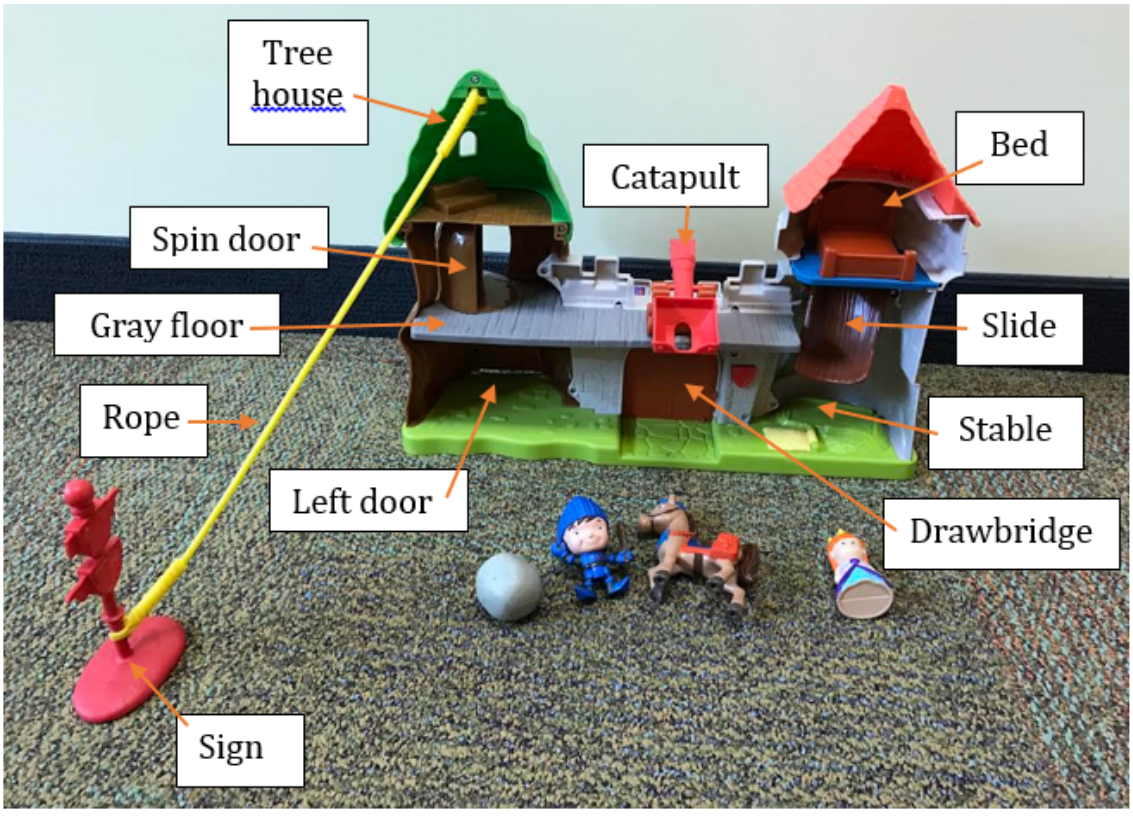

Figure 2. Example of play set with locations specified.

without the boulder in it or without using a figurine. In these examples, the play actions are missing movement, interaction with the play set, and a figurine respectively. We scored appropriate play actions as independent if the participant completed it without any assistance from the therapist. We also took data on prompted play actions, which we defined as appropriate play actions completed by the participant when the therapist provided any amount of physical prompt.

Independent appropriate play actions were considered different if they differed by one or more components from all other independently completed appropriate play actions within that session. For example, opening the castle door with the knight and making the knight and horse run were considered different play actions because they differed by action and location. Opening the castle door with the knight and opening the castle door with the princess were also considered different. Note that we calculated the number different by comparing independent play actions only. That is, we did not compare independent appropriate play actions to prompted appropriate play actions in order to calculate the number different. Thus, if a participant completed an independent play action that was the same as a previously prompted play action, we still counted this action different for that session. For example, if the therapist prompted the participant to make the knight open the castle door, and then later the participant independently made the knight open the castle door, this action was considered different. This was done in order to measure how many different play actions the participants themselves were completing each session.

Play actions were considered to meet the lag schedule in place if they were

different by one or more components from the previous number of play actions specified in the lag schedule. For the Lag 1 schedule (Leonard and Arthur), an independent play action needed to be different from the previous play action (whether independent or prompted). For the Lag 2 schedule (Jude), an independent play action needed to be different from the previous two play actions (whether independent or prompted). For example, the following string of 10 play actions includes eight actions that meet the Lag 1 schedule and six that meet the Lag 2 schedule ; opening the castle door with the knight (different), making the knight and horse run (different, meets Lag 1), making the princess and horse run (different, meets Lag 1 and Lag 2), opening the castle door with the princess (different, meets Lag 1 and Lag 2), making the knight and horse run (meets Lag 1 and Lag 2), making the knight and horse run, making the knight launch the boulder from the catapult (different, meets Lag 1 and Lag 2), making the princess and horse run (meets Lag 1 and Lag 2), opening the castle door with the knight (meets Lag 1 and Lag 2), and making the princess and horse run (meets Lag 1). That is, eight of the actions are different from the previous action by at least one component and six of the actions are different by at least once component from the previous two actions. In this example, the participant varied his responding across five different play actions.

The secondary dependent measure in this study was the number of cumulative novel independently completed appropriate play actions across sessions. We considered an appropriate play action to be novel if it was different by at least one component from all other appropriate play actions emitted previously in the study.

We recorded data using paper and pencil from video recordings of the sessions.

For each session, the data collector transcribed the play actions emitted by the participant in terms of figurine, movement, and location. The data collectors time stamped all play actions in order to facilitate analysis of agreement data. The data collectors also scored each response as independent or prompted, different (or not), and meeting the Lag 1 and/or Lag 2 schedule (or not). The data collectors then tallied the number of different play actions and number of actions meeting the lag schedule for each session.

We collected inter-observer agreement (IOA) data for an average of 30% of sessions across all conditions for each participant. IOA data were collected as described above. We calculated point-by-point IOA by comparing the lists of time stamped play actions from both data collectors. Agreements were scored if both data collectors recorded the exact same action and time stamp (within 3 s) and disagreements were scored if the two data collectors recorded different actions (different figurine, movement, or location) or the time stamp was off by more than 3 s. We then divided the number of agreements by the number of agreements plus disagreements and multiplied this number by 100 to yield a percentage. The IOA for each participant is listed in Table 4.

Design and Procedures

We evaluated the effects of multiple exemplar training plus discrimination training on generalization of variability of play behavior using a non-concurrent multiple baseline design. For Leonard and Arthur, we also embedded a multi-element design into the multiple baseline design. The order of phases was baseline, baseline probes, multiple exemplar training plus discrimination training (for Leonard and Arthur), multiple

Table 4

Interobserver Agreement and Treatment Integrity Data for Each Participant During Each Phase

	Leonard		Arthur		Jude	
Condition	IOA	Tx Integrity	IOA	Tx Integrity	IOA	Tx Integrity
Baseline/Baseline probes	83% (0-100%)	100%	86% (0-100%)	100%	94% (50-100%)	100%
Multiple Exemplar plus Discrimination	96% (80-100%)	98% (92-100%)	94% (84-100%)	97% (89-100%)	--	--
Multiple Exemplar Alone	93% (80-100%)	97% (85-100%)	84% (63-94%)	95% (87-100%)	88% (50-100%)	95% (91-100%)
Generalization	83% (63-94%)	97% (94-100%)	83% (50-100%)	91% (92-100%)	89% (76-100%)	99% (98-100%)

exemplar training alone (for all three participants), and generalization. Each phase and condition is described in detail below.

General Procedures

Prior to each session, the therapist conducted a one-trial multiple stimulus preference assessment to identify a preferred edible to deliver as reinforcement for that session. For the preference assessment, the therapist placed five edible items in a row in front of the participant and said, “Pick one.” The first item that was selected was used for that session. Sessions were 5 minutes in length. The therapist conducted an attending procedure by bringing the participant to the door of the research room and prompting him to remove the colored card that was attached to the door. The therapist then opened the door and prompted the participant to match the colored card to a colored bracelet hanging on the wall just inside the room, and helped the participant remove the bracelet from the wall and place it on his own wrist. The colored card and bracelet were designed to serve

as the discriminative stimuli corresponding to the contingencies in place for each session. The therapist then prompted the participant to the table that held the play set, said, “Go play,” and started the 5-minute timer. When the 5 minutes elapsed, the therapist said, “You’re all done” while pushing the play set out of reach and guided the participant out of the room.

Baseline

We conducted all baseline sessions as described under the general procedures. The colored card and bracelet during baseline sessions were yellow. The therapist did not provide any differential consequences for appropriate play actions and did not implement any prompting procedures. The purpose of baseline was to evaluate levels of variability prior to implementing contingencies of reinforcement that support variability.

Baseline Generalization Probes

Following the initial baseline, we conducted one to three sessions for each of the three generalization play sets to assess levels of play variability with these sets. We conducted all baseline generalization probes under baseline procedures, including the yellow card and bracelet.

Baseline Vary and Rep Probes

We conducted two sessions under baseline conditions except that the colored cards and bracelets were the colors that would be associated with variability and repetition during the multiple exemplar training plus discrimination training phase. We conducted one session with an orange card and bracelet, which would be associated with

repetition during the training phase, and one session with a purple card and bracelet, which would be associated with variability during the training phase. During the attending procedure, the colored card for the session was fixed to the door, and both bracelets hung on the wall just inside the room. The therapist prompted the participant to match the colored card below its respective bracelet, remove that bracelet from the wall, and place it on his own wrist. The play sets included in these probes was selected at random for each participant. The purpose of these probe sessions was to assess whether the discriminative stimuli that would be programmed during the multiple exemplar training plus discrimination training would have an effect on levels of variability prior to implementing the training.

Multiple Exemplar Plus Discrimination Training

Following the baseline generalization probes, we implemented the multiple exemplar training plus discrimination training with two of the participants (Leonard and Arthur). The purpose of this phase was to train the participants to engage in varied play behavior across multiple play sets as well as to establish a discriminative stimulus for variability and repetition. Throughout this phase, we alternated between sessions where variability was reinforced in the presence of a specific discriminative stimulus (Vary-purple card and bracelet) and sessions where repetition was reinforced in the presence of a different discriminative stimulus (Rep-orange card and bracelet). The procedures for variability and repetition sessions are described in detail below. Sessions also rotated across the three training play sets in such a way that the participants were exposed to each

play set an equal number of times in both the Vary and Rep conditions. We did this by conducting series of six sessions with the variability and repetition conditions occurring across each of the three play sets. Thus, we alternated across six types of sessions; (a) Vary 1, (b) Vary 2, (c) Vary 3, (d) Rep 1, (e) Rep 2, and (f) Rep 3 where Vary and Rep represent the variability and repetition conditions and the numbers represent the three play sets. We implemented the series of sessions in a semi-random order such that the participant was not exposed to the same condition (Vary or Rep) more than two sessions in a row.

Vary. We conducted the Vary sessions the same as described in the general procedures except that we reinforced appropriate play actions according to a Lag 1 schedule (Leonard and Arthur). We determined the lag for each participant based on their responding during baseline sessions. Leonard and Arthur's behavior rarely met a Lag 1 schedule during baseline, so we implemented the Lag 1 schedule during training and generalization with both participants. The purpose of this condition was to (a) teach the participant to vary his play behavior and (b) to establish the purple bracelet as a discriminative stimulus for variability that could then be used to promote generalization of variability to untrained play sets.

The therapist began each session by conducting the attending procedure with the purple card and bracelet, bringing the participant to the table with the play set, and saying, "Go play." The purple card and bracelet were intended to serve as the discriminative stimulus for variability. The therapist then waited 10 s for the participant to emit an appropriate play action. The therapist did not provide a consequence for the

first appropriate play action, whether prompted or independent. This allowed the participant to "set up" the lag schedules, such that he would have a response from which to differ. For the Lag 1 schedule of reinforcement (Leonard and Arthur), the therapist provided a small piece of edible following each appropriate play action that was different from the previous play action by one component (i.e., figurine, movement, or location).

The therapist implemented an error correction procedure if the participants (a) did not emit an independent appropriate play action for 10 s or (b) emitted a play action that did not meet the schedule of reinforcement (i.e., appropriate but did not meet schedule, and inappropriate play actions). For the error correction procedure, the therapist prompted the participant to complete three appropriate play actions that met the schedule of reinforcement. The therapist provided a small edible following each prompted response. The prompted play actions were selected by the therapist during the sessions (i.e., were not predetermined). If at any point during the error correction sequence, a participant completed play actions independently that met the schedule of reinforcement, the therapist discontinued the prompt sequence.

Repetition (Rep). We conducted the Rep sessions the same as described in the general procedures except that we reinforced appropriate play actions if they were the same as the first responses emitted that session. That is, we implemented a self-selected repetition contingency where we allowed the participant an opportunity to engage in any appropriate play action, which then became the only play action that would result in delivery of an edible for the remainder of that session. The play action was free to vary across repetition sessions, however. The purpose of the repetition condition was to serve

as a control for the variability condition.

The therapist began each session by conducting the attending procedure with the orange card and bracelet, bringing the participant to the table with the play set, and saying, “Go play.” The orange card and bracelet were intended to serve as the discriminative stimulus for repetition. The therapist then waited 10 s for the participant to emit an appropriate play action. The therapist did not provide a consequence for the first appropriate play action emitted by the participant. This allowed us to establish which response would result in delivery of the edible for that session. The therapist then provided an edible following all subsequent play actions if they were the same as the first play action. Similar to the Vary condition, the therapist implemented an error correction procedure if the participants (a) did not emit an appropriate play action for 10 s or (b) emitted appropriate play actions that were different from the first play action of the session. The error correction procedure was the same as in the Vary condition except the therapist prompted the participants to repeat the designated play action three times.

Prompt fading. After 20 sessions of exposure to the multiple exemplar training plus discrimination training, we did not observe increases in Leonard’s varied play behavior in the Vary condition over the Rep condition. Leonard was frequently contacting reinforcement for prompted behaviors during these sessions, but was emitting few independent responses that met the schedule of reinforcement. It appeared that he was becoming prompt dependent. Thus, at session 43 with Leonard, we made a modification to the way we prompted during error correction (denoted by the * on all graphs). From this session on, during the first error correction sequence of the session we

used full hand-over-hand prompts to prompt Leonard to engage in appropriate play actions that met the schedule of reinforcement (in both the Vary and Rep conditions). During all subsequent error correction sequences within a session, we used less intrusive physical prompts. The therapist initiated a prompted response by providing guidance from Leonard's forearm. If Leonard completed an appropriate play action that met the schedule of reinforcement, the response was reinforced. If he did not begin to complete an appropriate play action or began to move toward completing an action that did not meet the schedule, the therapist increased the intrusiveness of the prompt to Leonard's wrist, and then to Leonard's hand if he still did not complete an action that met the schedule of reinforcement. This prompt fading sequence was designed to be rapid and flexible, and took no more than two seconds to implement. Prompt fading was also included in the procedures for Arthur and Jude when they began the training phases. The purpose of making this change was to promote increases in independent responding by systematically fading the intrusiveness of our prompts within each session.

Social interaction plus differential reinforcement (Leonard). After implementing the prompt fading modification for Leonard, his independent appropriate play behavior decreased to zero across several sessions. Upon consultation, his clinical team informed us that Leonard tended to respond better to a combination of social plus edible reinforcement, opposed to just edible items alone. Therefore, at session 50 we added social interaction to the reinforcement procedure (denoted by the ** on all graphs). From session 50 on, prompted appropriate play actions that met the schedule of reinforcement resulted in delivery of an edible plus mild praise in a neutral tone of voice

(e.g., “Good job”) and independent appropriate play actions that met the schedule resulted in delivery of an edible plus enthusiastic praise (e.g., “Yeah!!! That’s it!”) and gentle physical touch (e.g., hair tousle, back rub, etc.). The purpose of this modification to the reinforcement procedure was to attempt to increase the effectiveness of our consequences as reinforcement as well as to differentially reinforce independent responding over prompted responding in an effort to reduce prompt dependence.

Increased wait time to 20 s, and stopped blocking inappropriate play behaviors. Following the two modifications described above for Leonard, we still did not observe increases in varied play behavior within the Vary component. Further, independent appropriate play actions were generally low for both Leonard and Arthur within both the Vary and Rep components. When we looked at their behavior within these sessions, it appeared that our error correction procedures were perhaps interfering with the participants completing independent play actions. That is, by blocking inappropriate play behavior and only allowing the participants 10 s before prompting, we were not allowing them to explore the play sets or try out different movements with the figures. Therefore, at session 68 for Leonard and 42 for Arthur, we made modifications to the criteria for implementing the error correction procedure (indicated by the # on all graphs). We extended the amount of time the participants had to engage in an appropriate play action from 10 s to 20 s. We also stopped implementing the error correction procedure following inappropriate play actions. This means that the participants were given 20 s to interact with the play set in any way that they chose. If they did not emit an appropriate play action within 20 s, the therapist conducted the three-prompt error correction

sequence as described above. If the participants did emit an appropriate play response but it did not meet the schedule of reinforcement, the therapist conducted the error correction procedure. These modifications were in place for Jude when he began the multiple exemplar training.

Multiple exemplar training alone. After 56 sessions of exposure to the discrimination training procedures for Leonard and 27 sessions for Arthur (and multiple procedural modifications), it appeared that the participants were not acquiring a strong discrimination between repetition and variability. Further, the participants were not emitting enough independent appropriate play actions to repeatedly contact the contingencies in place for the two components of the discrimination training. We hypothesized that the participants would be more successful at contacting the contingencies for variability, and thus would begin to emit more independent and varied appropriate play actions, if they had repeated and continuous exposure to just the variability contingencies. Thus, at session 77 for Leonard and 56 for Arthur, we discontinued the discrimination training component of this study and began implementing Vary sessions only. This change was put in place for Jude at session 43, after we ran only one Rep session with him.

All sessions from this point forward were conducted as described above, except that we rotated across the three training play sets using the Vary condition only. Leonard's and Arthur's independent appropriate play behavior were reinforced on a Lag 1 schedule of reinforcement as described above. Jude's independent appropriate play behavior was reinforced on a Lag 2 schedule of reinforcement. This was determined

based off of his behavior during baseline. Because Jude's play behavior frequently met a Lag 1 schedule but not a Lag 2 schedule during baseline, we decided to implement a Lag 2 schedule during the training and generalization phases of the study for him. As described above, the therapist did not provide a consequence for the first responses in each session. For the Lag 2 schedule of reinforcement, the therapist provided a small edible following the second response of each session if it was different from the first, and the Lag 2 schedule took effect on the third response. That is, from the third response on, the therapist provided an edible if the appropriate play action was different by one component from the previous two play actions.

Prompt to action (Leonard only). When we discontinued the discrimination training component of the study, we saw an increase in independent varied play behavior for Leonard with the farm play set, but not with the other two play sets (castle and fire station). Anecdotally, Leonard was still showing signs of prompt dependency. For example, he would frequently pick up a figuring and then play his hand in the therapist's hand. Thus, at session 100 for Leonard, we introduced a final change to the prompting procedure (indicated by the ## on all graphs). Following 20 s of no independent appropriate play actions, the therapist prompted Leonard's hand towards the play set but did not prompt him to actually complete any play actions. This entailed the therapist nudging Leonard's elbow in such a way that his arm and hand moved towards the set. If Leonard then completed an independent play action that met the Lag 1 schedule, the therapist treated this as an independent action and provided an edible plus enthusiastic praise and gentle physical touch. If Leonard did not complete an independent play action

within 5 s of the prompt to action, the therapist provided another prompt to action. This was repeated up to three times, at which point the therapist implemented the full error correction procedure. The purpose of this modification was to reduce prompt dependency and promote independent responding such that Leonard's behavior would contact the differential reinforcement contingency.

Social interaction plus differential reinforcement (Jude). Following 12 sessions of exposure to the multiple exemplar training alone, we did not observe increases in Jude's independent varied play behavior. Upon consultation with his clinical team, we learned that Jude, similar to Leonard, had a history of responding to social praise plus edibles, instead of just edibles alone. Thus, at session 55 for Jude we added social interaction to the reinforcement procedures for Jude. We implemented the same modification made to Leonard's procedures there were described above. This change is also indicated by the ** on all graphs.

Modifications to reinforcement procedures (Jude only). Beginning at session 61, we made a series of modifications to the reinforcement procedures for Jude. Following the introduction of differential social reinforcement, Jude continued to emit very few independent play actions. We hypothesized that the edibles plus social praise were not functioning as reinforcement for him. At sessions 61 (denoted by the 1 on all graphs) we included a new item (frosting) into the preference assessment that we conducted prior to each session, and removed an item that he had chosen the fewest number of times (chocolate chip). At session 62 (denoted by the 2 on all graphs), we made access to edibles contingent on independent responding only. That is, prompted responses resulted

in neutral praise only and independent responses resulted in an edible plus enthusiastic praise plus gentle touch. At session 64 (denoted by the 3 on all graphs), we began to offer Jude a choice of edibles following each independent response. Rather than conducting the five-item preference assessment prior to the session, we allowed him to choose from the five items after each independent response (by offering him a small plate with the edibles on it). This procedure took approximately 2 seconds to complete and did not interfere with the session time. The purpose of these modification was to increase the reinforcing effectiveness of our programmed consequences for independent varied behavior.

Pre-exposure (Jude only). At session 66 (denoted by the 4 on all graphs), we began to implement a pre-exposure procedure where the therapist would prompt Jude to complete three appropriate play actions that were all different from each other. The therapist provided access to the plate of edibles, plus neutral praise, for the second and third prompted actions. Following the third prompted action, the therapist started the five-minute timer and began the session. For the remainder of the session, only independent appropriate play actions that met the Lag 2 schedule resulted in access to the plate of edibles. All other procedures during the sessions were as described above. The purpose of this procedure was to bring Jude's behavior in contact with the contingency of "different play actions equals treats" while also reducing prompt dependence throughout the session.

Generalization. Following multiple exemplar training alone, we assessed generalization of variability of independent appropriate play actions to three untrained play sets for each participant. For the purpose of this study, we defined generalization as

the rapid acquisition of variability with untrained play sets. That is, we considered varied play behavior to have generalized to the untrained play sets if it was observed to occur at similar levels as were observed in the previous treatment phase within three to five sessions of exposure to each play set. Thus, the Vary reinforcement contingencies (Lag 1 or Lag 2) were in place, rather than testing for generalization under baseline conditions. For each participant, we tested for generalization to one generalization play set at a time (refer to Table 3 for the lists of play sets for each participant). For the first generalization play set, we conducted sessions where the Lag 1 (Leonard and Arthur) or Lag 2 (Jude) contingencies were in place, but the error correction procedures were not. During these sessions, all modifications made to the reinforcement procedures were in place for each participant, but none of the modifications made to the prompting procedures or error correction were in place (see Table 5 for a summary of procedural modifications). That is, Leonard's independent appropriate play actions resulted in an edible plus enthusiastic praise if they met the reinforcement schedule, Arthur's play actions resulted in an edible only, and Jude's play actions resulted in access to the plate of edibles plus enthusiastic praise. If generalization of varied play behavior was observed, then we tested generalization to the next play set. If generalization was not observed, we implemented the error correction procedure with that play set (including all modifications made to the prompting and error correction). Once independent varied play actions increased with the play set, we tested for generalization to the next play set (sans error correction). For Leonard, we implemented the generalization procedures for two of the generalization play sets, and implemented the error correction procedure for the third generalization

Table 5

Summary of Procedural Modifications

	Modification (session number/symbol)	
Participant	MET plus DT	MET Alone
Leonard	Prompt fading (43/*)	Prompt to action (100/##)
	Differential reinforcement for social interaction (50/**)	
	Increase wait time to 20 s (68/#)	
Arthur	Increase wait time to 20 s (42/#)	NA
Jude	NA	Differential reinforcement for social interaction (55/**)
		Added frosting (61/1)
		Differential reinforcement of
		edibles (62/2)
		Choice of edibles for independent responding (64/3)
		Pre-exposure (66/4)

Note. MET = multiple exemplar training; DT = discrimination training.

play set. For Arthur, we implemented the error correction procedure for the first play set, and implemented the generalization procedure for the remaining two play sets. For Jude we implemented the generalization procedures (without error correction) for all three generalization play sets.

Treatment integrity. We assessed integrity of implementation of the procedures for an average of 28% of sessions across all conditions for all participants. For these sessions, we scored whether or not the therapist correctly conducted the preference assessment, conducted the attending procedure, prepared the play set, and gave the correct instruction to begin the session. We also scored per opportunity treatment

integrity throughout the sessions on whether the therapist implemented the correct consequence for each of the participants' responses. For each appropriate play action (prompted and independent), we scored whether the therapist delivered reinforcement (for responses that met the contingency) or correctly implemented the error correction procedure (for responses that did not meet the contingency, or for non-responding during the wait period). As modifications were made to the reinforcement, prompting, and error correction procedures, these components were also scored as being implemented correctly or incorrectly by the therapist. Each of these items was scored as a "yes," "no," or "NA." We calculated the total treatment integrity score for each session by dividing the number of therapist behaviors implemented correctly divided by the total number of opportunities for the therapist to implement the procedures and multiplying by 100 to yield a percentage. The treatment integrity scores for each participant can be found in Table 4.

CHAPTER IV

RESULTS

For all graphs, Leonard's data are displayed in the top panel, Arthur's in the middle panel, and Jude's in the bottom panel. The x-axis represents sessions. Each shape represents a different play set, where closed shapes represent sessions conducted under the Vary condition and open shapes represent sessions conducted under the Rep condition. Refer to Table 3 for a list of which play sets were assigned to which condition, and which symbols represent which play sets for each participant.

Number of Play Actions Meeting Lag Schedule

Figure 3 represents the number of independently completed appropriate play actions that met the Lag 1 schedule (Leonard and Arthur) and Lag 2 schedule (Jude).

Leonard

During baseline, Leonard completed up to two actions that met the Lag 1 with the fire station, up to one action with the castle, and up to seven actions with the farm play set. During the baseline probes, he completed seven actions that met the Lag 1 schedule with the pirate ship play set, but then did not complete any play actions that met the Lag 1 during the next two sessions with the pirate ship. He completed one play action that met the Lag 1 with the Lion King set, and did not complete any actions that met the Lag 1 with the Paw Patrol set or during the Vary and Rep probes (with the castle).

Upon implementation of the multiple exemplar training plus discrimination

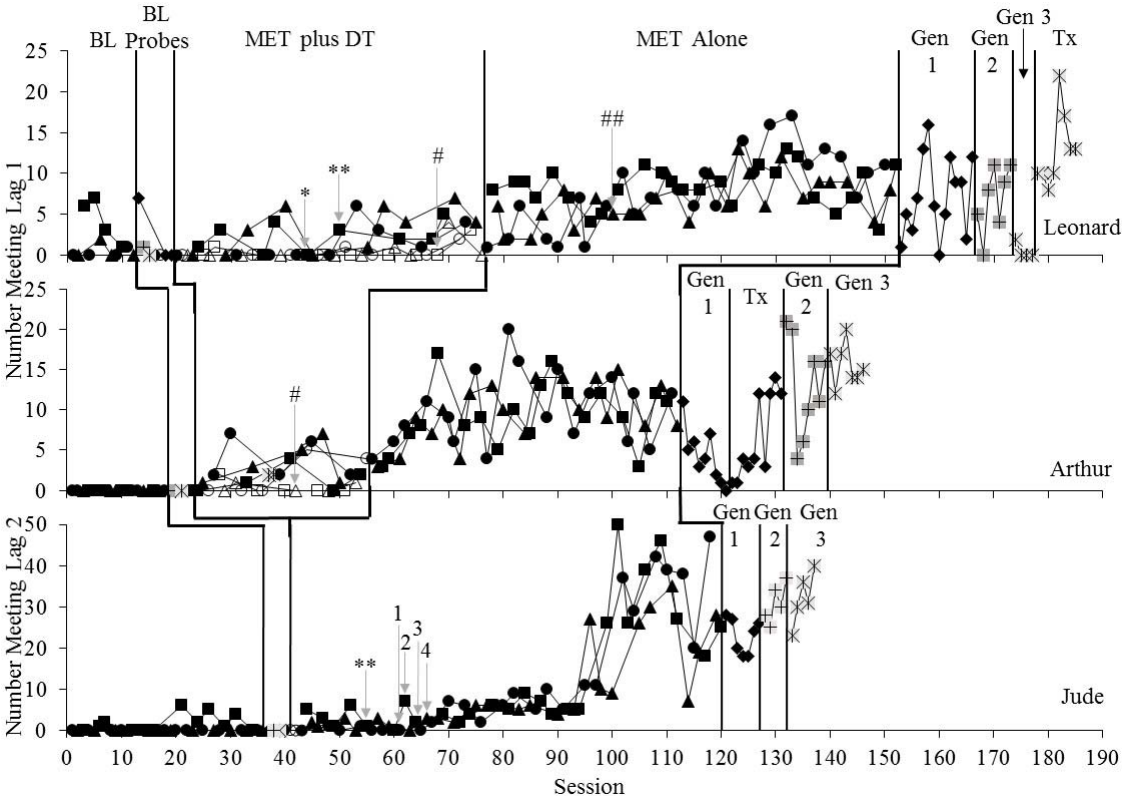

Figure 3. Number of independent appropriate play actions meeting the Lag 1 schedule for Leonard and Arthur, and meeting the Lag 2 schedule for Jude. BL = baseline; MET plus DT = multiple exemplar training plus discrimination training; MET alone = multiple exemplar training alone, Gen Set = generalization set; Tx = training.

training, we did not observe large increases in play actions that met the Lag 1 schedule in either the Vary or Rep components. We did see slight increases in the Vary sessions, but none above seven play actions meeting the Lag 1. In general (and anecdotally), we noticed that Leonard was showing signs of prompt dependency (e.g., placing his hand in the therapist's hand). We also hypothesized that edibles were not functioning as reinforcers for him in these sessions. Upon consultation, Leonard's clinical team informed us that he did have a tendency to be prompt dependent during teaching sessions

and that he responded well to social interactions as reinforcement. Thus, we made a series of modifications to target Leonard's tendency towards prompt dependence and to increase the efficacy of our programmed consequences as reinforcers. At session 43 we introduced the prompt fading modification (indicated by the * on the graph). Following seven sessions of exposure to this modification, the number of play actions meeting the Lag 1 schedule decreased to 0. At session 50 we introduced social interaction delivered differentially for independent play actions that met the Lag 1 contingency (indicated by the **). Over the next 18 sessions, we observed increases in the number of play actions meeting the Lag 1 during the Vary component, but these increases were not above baseline levels. Anecdotally, it appeared that our criteria for when to prompt were interfering with Leonard completing independent play actions. That is, by only allowing him 10 s to interact with the play set prior to conducting the error correction procedure and by blocking all inappropriate play, it seemed that we were preventing him from exploring the play set and completing play actions on his own. Thus, at session 68 we extended the wait time to 20 s and we also stopped blocking inappropriate play behaviors (indicated by the #).

During the next nine sessions, we observed increases in the number of play actions meeting the Lag 1 during the Rep component. It appeared that Leonard's behavior was not coming under control of the discrimination between variability and repetition. In addition, we were not observing large increases in variability of play actions during the Vary component. We therefore decided to discontinue the discrimination portion of the study for Leonard at session 77, and began to implement multiple exemplar training alone

(i.e., Vary only). Over the next 32 sessions, we observed some increases in the number of play actions meeting the Lag 1, primarily with the farm play set. During these sessions, it appeared that Leonard was still dependent on our prompts. For example, he frequently picked up a figurine and then attempted to place his hand into the therapist's hand and wait for her to prompt him. At session 100 we introduced the prompt to action, in an attempt to further reduce prompt dependency (indicated by the ##). Following this change, we saw steady increases in the number of independent play actions meeting the Lag 1 schedule with all three play sets. Leonard completed up to 17 play actions that met the Lag 1 with the castle set and up to 13 play actions that met the Lag 1 with the fire station and farm sets. Although we observed a high degree of variability in the number of play actions meeting the Lag 1 schedule, Leonard's behavior fell within a range between three play actions and 13 play actions that met the Lag 1 schedule between sessions 139 and 152. We therefore determined that Leonard was demonstrating varied play behavior with these three play sets under the training conditions.

Following the multiple exemplar training alone phase, we tested for generalization of Leonard's varied play behavior to the first generalization play set (pirate ship). Initially, Leonard completed one play action that met the Lag 1 schedule. His behavior then increased over the next several sessions until session 158 where he completed 16 play actions that met the Lag 1 schedule. His behavior was then variable between zero and 12 play actions that met the Lag 1 schedule. Although Leonard's data were variable, during the majority of sessions he completed between three and 12 play actions that met the Lag 1 schedule. Thus, it appeared that his varied play behavior did

generalize to the pirate ship play set. We saw a similar pattern of responding with the second generalization play set (Lion King), with the majority of independent appropriate play actions meeting the Lag 1 schedule between three and 11 per session. When we introduced the third generalization play set (Paw Patrol), Leonard completed two play actions that met the Lag 1 schedule, and then zero play actions that met the Lag 1 over the next three sessions. We then introduced the error correction procedure for this play set at session 178 and saw an immediate increase in the number of play actions that met the Lag 1 schedule to 10. At session 182, this increased to 22 play actions that met the Lag 1 schedule, and then decreased over the next few sessions to 13.

Arthur

During baseline, and all baseline probes, Arthur did not complete any play actions that met the Lag 1 schedule. Upon implementation of the multiple exemplar training plus discrimination training, we saw increases in the number of play actions meeting the Lag 1 in all three play sets during the Vary components. We also saw some increases in the number of play actions meeting the Lag 1 during the Rep components for two play sets (Lion King and Doc McStuffins). At sessions 37 and 38, generalization set three was introduced to the training procedures under both the Vary and Rep components. This was done in error, and occurred for those two sessions only. At session 42 we increased the wait time from 10 s to 20 s, and also stopped blocking inappropriate play behaviors during the wait time (as indicated by the # on the graph). This was the same change we made to Leonard's procedures at session 68. Following this change we saw increases in the number of play actions meeting the Lag 1 up to seven at session 47. Over the next

seven sessions, the number meeting the Lag 1 schedule in the Vary component decreased to about two play action whereas the number meeting the Lag 1 in the Rep component was elevated to four play actions. It appeared that Arthur was not acquiring the discrimination between the two components and we were not seeing large increases in the number of play actions meeting the Lag 1 within the Vary component.

Therefore, at session 56 we discontinued the discrimination training component of the procedures with Arthur, and began implementing multiple exemplar training alone (i.e., Vary only). Following this change, we saw a steady increase in the number of play actions meeting the Lag 1 schedule across all three play sets. At session 81 Arthur completed 20 play actions that met the Lag 1 schedule (with the Lion King set). Following session 81, Arthur's behavior was variable between three and 16 play actions meeting the Lag 1 schedule. However, the majority of sessions fell between seven and 12 play actions meeting the Lag 1.

Following the multiple exemplar training alone, we tested for generalization of Arthur's behavior to the first generalization play set (house). Initially, Arthur completed 12 play actions that met the Lag 1 schedule. His behavior then decreased over the next eight sessions until session 121 where he did not complete any play actions that met the Lag 1 schedule. At session 122 we introduced the error correction procedure with the house play set, and saw increases in the number of play actions that met the Lag 1 to 12 play actions that met the Lag 1 schedule. His behavior then decreased to three play actions at session 128, and increased again to 12 and 13 play actions during sessions 129 through 131. When we tested the next generalization play set (farm) Arthur completed

between four and 27 play actions that met the Lag 1 schedule, with the majority of sessions above 10 play actions. When we tested for generalization to the final play set (Paw Patrol), Arthur completed between 12 and 20 play actions that met the Lag 1 schedule. The error correction procedure was not required for the second and third generalization play sets, thus Arthur's behavior generalized to these two play sets. However, it is important to remember that the third play set (Paw Patrol) was briefly exposed to the training procedures earlier in the study at sessions 37 and 38, thereby weakening conclusions to be drawn regarding generalization of Arthur's behavior to this set.

Jude

During baseline for Jude, he did not complete any play actions that met a Lag 1 or Lag 2 schedule with the Paw Patrol and Farm play sets. With the pirate play set, he completed up to 43 play actions that met a Lag 1 schedule. However, upon inspection of the raw data, it was clear that Jude was engaging in higher order stereotypy by alternating between the same two play actions for the majority of those sessions. When these data were analyzed under a Lag 2 schedule, we saw that up to six of his play actions met a Lag 2 schedule. Therefore, we analyzed Jude's data in terms of a Lag 2 schedule, and implemented a Lag 2 schedule of reinforcement during training and generalization. During the baseline probes, Jude did not complete any play actions that met the Lag 2 schedule. We implemented one Rep session (session 42) and then implemented multiple exemplar training only (i.e., Vary only). Given the failure to achieve discriminated responding with the other two participants, we chose to implement only the multiple

exemplar training alone phase with Jude.

Upon implementation of the multiple exemplar training alone, we did not see increases in Jude's responding above baseline levels. We hypothesized that edibles alone were not functioning as reinforcement during these sessions. Upon consultation, Jude's clinical team reported that edibles were not consistently effective as reinforcers during teaching sessions and that he responded to social interaction. Therefore, at session 55 we introduced social interaction that was differentially delivered for independent play actions that met the Lag 2 schedule (indicated by the **). During the following six sessions we did not observe increases in Jude's behavior. We then implemented a series of procedural modifications in an attempt to increase the effectiveness of our programmed consequences as reinforcement for independent play actions that met the Lag 2 schedule. At session 61, we added frosting to the pre-session preference assessment (and removed chocolate chip, indicated by the 1). Although he selected the frosting during the preference assessment for that session, this did not seem to affect his responding within the session. At session 62 we began to only deliver the selected edible for independent responses that met the Lag 2 schedule (indicated by the 2). During this session Jude completed seven play actions that met the Lag 2 schedule (the highest during the treatment phase up to this point). At session 63, however, he did not complete any play actions that met the Lag 2 schedule. At session 64 we discontinued the pre-session preference assessment and began to present a choice of edibles following each independent response (indicated by the 3). During sessions 64 and 65 Jude completed two and zero play actions that met the Lag 2, respectively. Anecdotally, during these two

sessions Jude reached for the plate of edibles multiple times and showed signs of distress when he was denied access to the edibles. It appeared that there was motivation to obtain them, but he was not completing play actions independently that would result in access to the edibles.

At session 66, we introduced the pre-exposure procedure. This allowed the therapist to expose Jude to the contingency (i.e., varied play actions results in access to the plate of edibles) while maintaining differential reinforcement for edibles within the sessions. Following this modification, we saw a slow and steady increase in the number of play actions that met the Lag 2 schedule, up to 10 actions at session 88. At session 96, we saw a sharp increase in the number of Jude's play actions that met the Lag 2 schedule to 27. This increase continued across all three play sets, up to 50 play actions during session 101. Over the next 19 sessions, Jude completed between seven and 46 independent play actions that met the Lag 2 schedule, with the majority of sessions falling between 18 and 40 play actions meeting the Lag 2 per session.

When we tested for generalization of Jude's varied play behavior to the first generalization play set (Lion King), he completed between 16 and 29 independent appropriate play actions that met the Lag 2 schedule. When we tested the next two generalization play sets, Doc McStuffins and Mini Mouse, Jude completed between 25 and 34 play actions that met the Lag 2 schedule with the Doc McStuffins play set and between 29 and 40 play actions that met the Lag 2 schedule with the Mini Mouse play set. Jude's varied play behavior generalized to all three generalization play sets in the absence of the error correction procedure.

Number of Different Play Actions

Figure 4 represents the number of different play actions completed independently for each participant. In general, Leonard and Arthur's patterns of responding were similar to that described for the number of independently completed play actions that met the Lag 1 schedule. That is, we observed increases and decreases in this measure that correspond to increases and decreases in the number of play actions meeting the Lag 1.

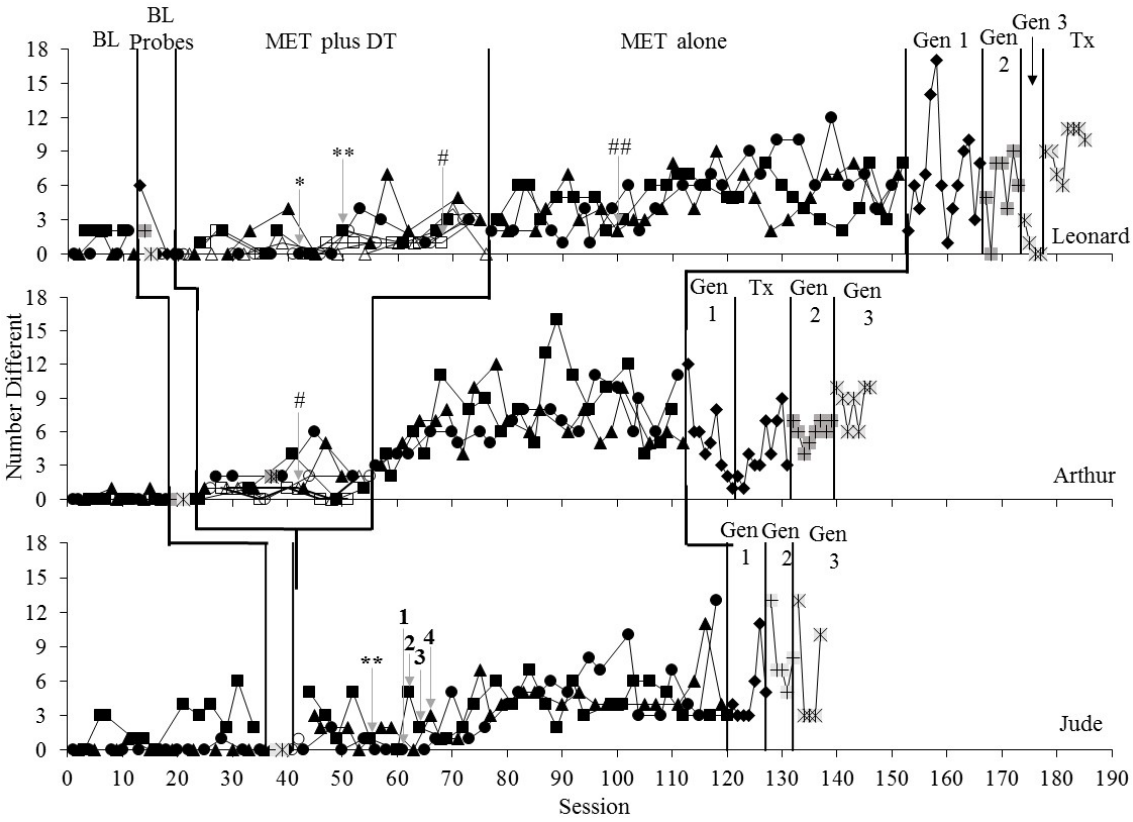

Figure 4. Number of different independent play actions for all participants. BL = baseline; MET plus DT = multiple exemplar training plus discrimination training; MET alone = multiple exemplar training alone, Gen Set = generalization set; Tx = training

Leonard

During baseline, Leonard completed up to two different play actions with the farm and castle play sets, and no actions with the fire station play set. During the baseline probes, he completed seven different play actions with the pirate play set, which then decreased to zero over the next two sessions. He also complete two different actions with the Lion King set. All other baseline probe sessions were at zero. Following introduction of the multiple exemplar training plus discrimination training, we observed slight increases in the number of different play actions primarily in the Vary component (up to seven). However, for a majority of Vary sessions in this phase Leonard completed between zero and five different play actions. We observed increases in the number of different play actions within the Rep component (up to four) after increasing the wait time to 20 s and discontinuing blocking inappropriate play actions (indicated by the #, around session 70).

When we discontinued the discrimination training with Leonard and began the multiple exemplar training alone (Vary only), we did not observe an increase in the number of different play actions above the highest point in the previous phase. We did, however, observe increases in the lower range of the number different. That is, during the multiple exemplar training plus discrimination training phase, Leonard completed zero or one different play action during many of the sessions. When we began the multiple exemplar training alone phase, Leonard always completed at least two different play actions, and usually he completed three or more. After introducing the "prompt to action" (indicated by the ##) we observed a steady increase in the number of different play

actions with all three play sets. Around session 123, we began to observe variability in Leonard's data, with the number of different play actions completed each session varied between three and 12 across the three play sets. We observed a sharp decrease in the number different for two play sets (fire station and farm) whereas we observed increases with the castle set. The number of different play actions completed with the fire station and farm sets increased around sessions 138 and 146, respectively, such that they were around the same level as the number different for the castle play set.

When we tested for generalization of Leonard's behavior to the first generalization play set (pirate ship), he initially completed two different play actions. The number of different play actions then increased sharply over the next five sessions to 17 different actions at session 158. Leonard's behavior then decreased sharply to one play action at session 160, but then increased again to between three and 10 different play actions across the last several sessions with the pirate play set. Although Leonard's behavior was variable during with this play set, on most sessions he completed between three and 10 play actions, providing further evidence that Leonard's varied play behavior generalized to this play set. When we tested for generalization to the next play set, Lion King, Leonard completed between four and nine different play actions per session (with the exception of session 168 where he did not complete any play actions). Therefore, we determined that his behavior had also generalized to the Lion King play set. When we tested for generalization to the final play set, Paw Patrol, Leonard completed three different actions, and then this decreased to zero over the next three sessions. It did not appear that Leonard's behavior generalized to this play set. When we implemented the

error correction procedure with the Paw Patrol play set, Leonard completed between six and 11 different play actions per session.

Arthur

Arthur completed up to one different play action during baseline, and completed zero play actions during the baseline probes. When we introduced the multiple exemplar training plus discrimination training, we observed slight increases in the number of different play actions completed up to two per session in the Vary component and one per session in the Rep component. At session 41, we saw further increases in the number of different play actions completed in the Vary component (up to six play actions in session 45). The number of different play actions that Arthur completed independently then decreased across all three play sets to around two play actions per session around session 53. When we discontinued the discrimination training and began the multiple exemplar training only phase, we saw an immediate and steady increase in the number of different play actions Arthur completed each session across all three play sets. From about session 72 to 111, Arthur completed between 4 and 16 different play actions each session, with the majority of sessions between 5 and 12 play actions per session.

When we tested for generalization of Arthur's behavior to the first generalization play set (house), he initially completed 12 different play actions. The number of different play actions completed pre session then decreased steadily over the next several sessions to one play action at session 121. We determined that Arthur's behavior did not generalize to this play set. When we introduced the error correction procedures with the house play set, the number of different play actions that Arthur independently completed

increased steadily to 10 different play actions at session 140. At session 141, the number of different play actions decreased to three. However, because Arthur was completing up to 12 play actions that met the Lag 1 schedule, we concluded that he had acquired variability with the house play set. When we introduced the next generalization play set (farm), Arthur completed between four and seven different play actions. We concluded that the number of different play actions Arthur completed independently generalized to the house play set. When we introduced the final generalization set, Paw Patrol, we saw a similar pattern of responding (recall that Paw Patrol was exposed to the training procedures at sessions 37 and 38).

Jude

During baseline, Jude did not complete any play actions with the Paw Patrol play set, completed one play action with the farm play set, and completed up to six different play actions with the pirate play set. Although he completed more than two different play actions with the pirate set during several baseline sessions, inspection of the raw data shows that he primarily alternated between two different play actions during these sessions (i.e., engaged in higher-order stereotypy). Jude did not complete any play actions during the baseline probes. At session 43 we conducted one session under the Rep condition, in which Jude independently completed one different play action. When we began the multiple exemplar training alone (Vary only) phase at session 44 we saw slight increases in the number of different play actions that Jude was completing with the Paw Patrol and farm play sets, but did not observe increases in the number of different play actions completed with the pirate set above baseline levels. From session 44 to about

session 66, Jude completed between zero and three different play actions with the Paw Patrol and farm play sets, and between zero and five different play actions with the pirate play set.

Following the series of procedural modifications for Jude (indicated by the **, 1, 2, 3, 4) we observed a slow and steady increase in the number of different play actions completed independently by Jude. These increases were seen primarily with the Paw Patrol and farm play sets. The number of different play actions did not increase above baseline levels for the pirate play set. From about session 181 to 120, Jude competed between four and 11 different play actions across all three play sets, with the majority of sessions between four and eight play actions. When we tested for generalization to the first generalization play set (Lion King), Jude completed between three and 11 different play actions per sessions. With the second set, Doc Mc Stuffins, he completed between five and 13 different play actions. With the third generalization play set, Mini Mouse, he completed between three and 13 different play actions.

Number of Cumulative Novel Play Actions

Figure 5 represents the cumulative novel play actions completed across the entire study for each participant. Table 6 displays the total number of novel play actions that each participant completed for each play set.

Leonard

For Leonard, in general, the slope of novel play actions was steeper during the multiple exemplar training alone phase than during the multiple exemplar training plus

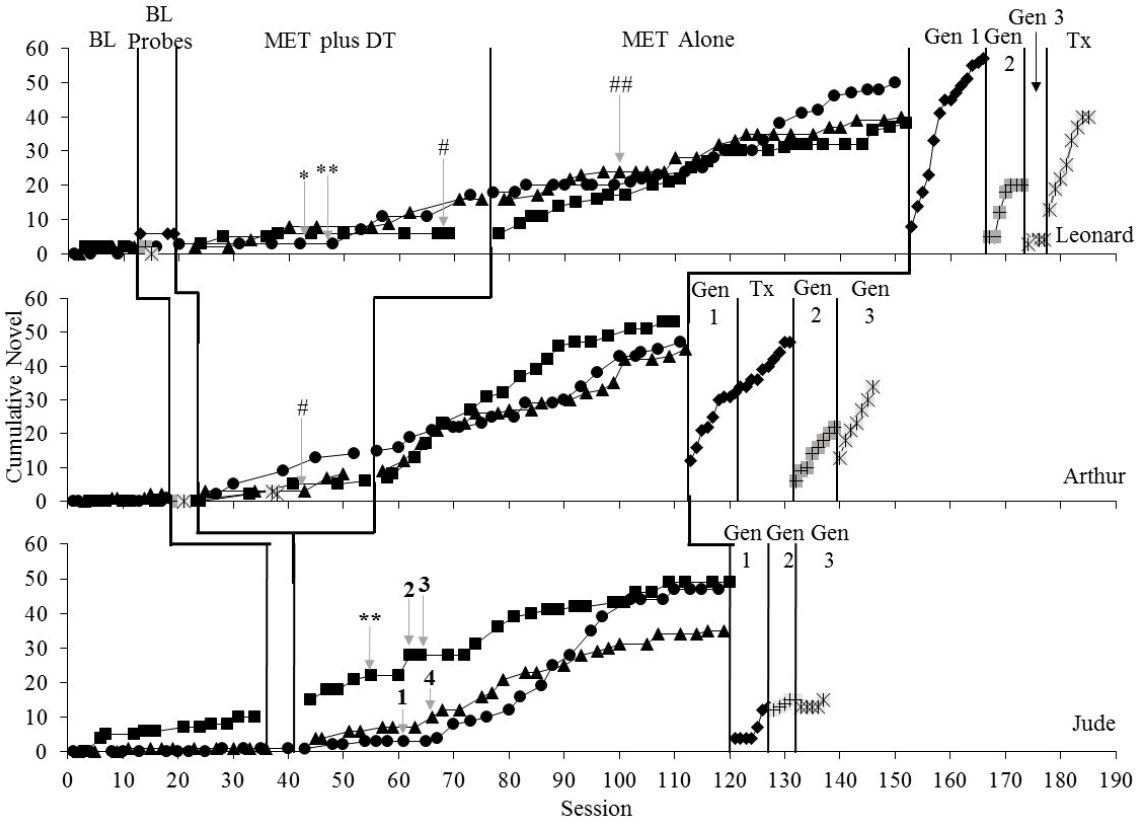

Figure 5. Cumulative novel play actions for each participant. BL = baseline; MET plus DT = multiple exemplar training plus discrimination training; MET alone = multiple exemplar training alone, Gen Set = generalization set; Tx = training.

Table 6

Total Novel Play Actions Completed for Each Play Set for Each Participant

	Leonard		Arthur		Jude	
Condition	Play set	Total	Play set	Total	Play set	Total
Training	Castle	50	Castle	45	Paw Patrol	35
	Fire Station	38	Doc McStuffins	53	Farm	47
	Farm	40	Lion King	47	Pirate Ship	49
Generalization	Paw Patrol	57	Farm	47	Lion King	13
	Lion King	20	Paw Patrol	22	Doc McStuffins	15
	Pirate Ship	40	House	34	Mini Mouse	15

discrimination training phase. Around session 116, the slope for novel play actions completed with the fire station and farm sets became more level, whereas the slope for the castle play set remained about the same. In general, the data from the two training phases indicates that Leonard was completing more novel play actions under the multiple exemplar training alone condition. The slope of novel play actions with the first generalization play set (pirate ship) was steeper than during the training phase, and he completed more novel play actions during this phase than during the training phase. The slope with the second generalization play set (Lion King) was initially steeper than during the training phase, but then leveled off during the last three sessions with this set. The slope with the third generalization play set (Paw Patrol) was initially zero, but when we introduced the error correction for this play set the slope increased to a similar slope as with the first generalization play set.

Arthur

In general, for Arthur, the slope of novel play actions was steeper across all three play sets in the multiple exemplar training alone phase than during the multiple exemplar training plus discrimination training phase. This indicates that Arthur was completing more novel play actions during the multiple exemplar training alone phase. The slope with the first generalization play set (house) was initially steeper than during the training phase, but then began to level out until we introduced the error correction procedure with this play set. The slopes for the second and third generalization play sets were also steeper than during the multiple exemplar training alone phase.

Jude

For Jude, there was a slight slope for the pirate ship play set during baseline, but no slope for the other two play sets. During multiple exemplar training alone, the slope of novel play actions with the pirate ship play set increased slightly. The slope for the other two play sets (farm and Paw Patrol) did not begin to increase until following the series of procedural modifications around session 66. The slope for the Paw Patrol play set (triangles) increased to about the same slope as pirate ship (squares). The slope for farm (circles) increased around session 66, but then continued to become steeper until the total number of novel play actions for this play set was at the same level as for the pirate ship set. The slope of novel play actions with the first generalization play set (Lion King) was initially zero, but then increased over the last three sessions with this play set. With the remaining two generalization play sets, Doc McStuffins and Mini Mouse, he completed 12 and 13 different play actions respectively, but then completed very few novel play actions. Thus, the slope for the final two generalization play sets was low.

Number of Independent Appropriate Play Actions

In addition to measuring variability of play behavior, we also measured the general level of appropriate play behaviors that the participants engaged in. Figure 6 displays the number of independent appropriate play actions per session for each participant. In general, we found that the number of independent appropriate play actions completed each session roughly corresponded to the number of play actions that met the Lag 1 (Leonard and Arthur) and Lag 2 (Jude) schedules.

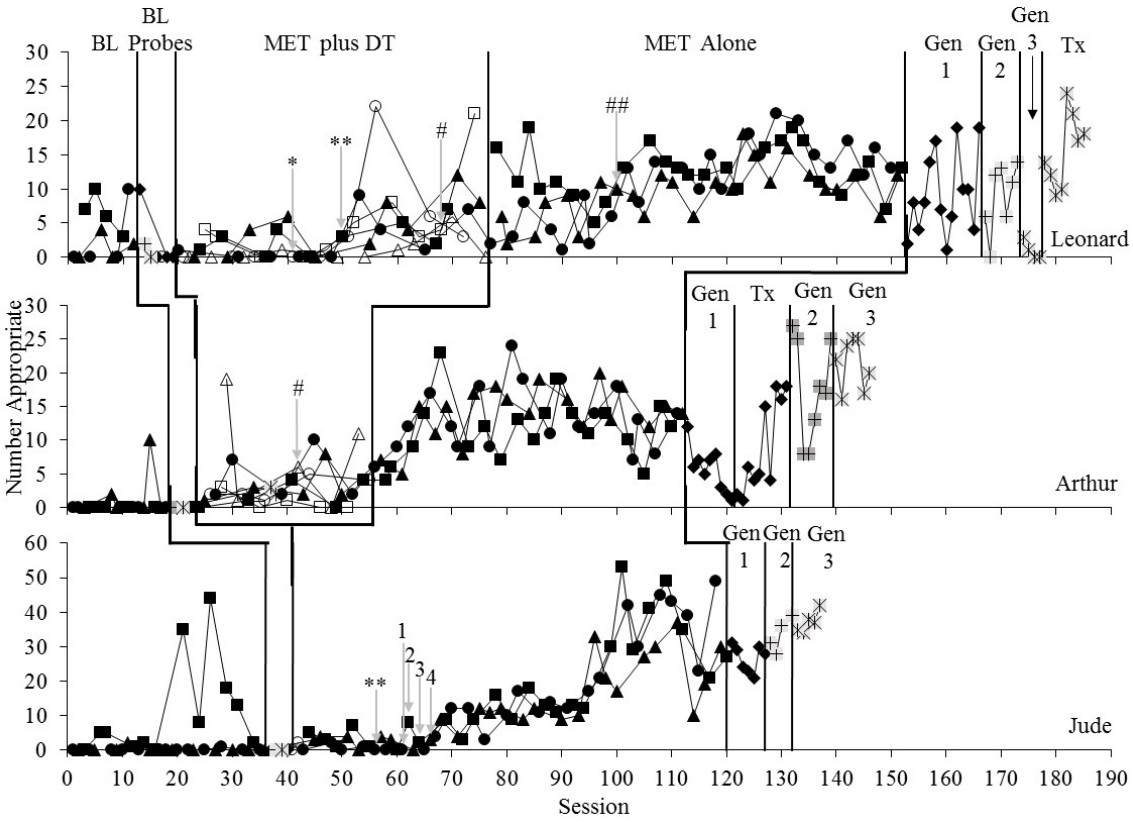

Figure 6. The number of independent appropriate play actions per session for each participant. BL = baseline; MET plus DT = multiple exemplar training plus discrimination training; MET alone = multiple exemplar training alone, Gen Set = generalization set; Tx = training

Leonard

During baseline, Leonard completed up to 10 independent play actions, and he completed 10 play actions during one of the baseline probe sessions. During the multiple exemplar training plus discrimination training phased, we did not observe increases in independent appropriate play actions until after introducing differential social reinforcement. However, these increases were still not higher than baseline levels. When we implemented the multiple exemplar training alone phase, we saw increases in Leonard's appropriate play actions with the farm play set, but not with the castle and fire

stations sets. After introducing the "prompt to action" we saw increases in Leonard's independent appropriate play actions across all three play sets. With the first two generalization play sets (pirate ship and Lion King), we did not introduce error correction, and Leonard completed up to 19 and 14 independent play actions per session, respectively. Leonard completed four independent play actions with the final generalization play set (Paw Patrol). When we introduced error correction for this set, Leonard's independent play actions increased to above levels seen during the multiple exemplar training alone phase.

Arthur

Arthur did not complete any independent play actions during baseline, with the exception of sessions 8 and 15 where he completed up to 10 independent actions with the castle play set. When we began the multiple exemplar training plus discrimination training phase we observed increases in the number of appropriate play actions up to 10 actions per session. When we implemented the multiple exemplar training alone phase, Arthur's independent appropriate play actions increased steadily. For the majority of this phase, he completed between 12 and 19 independent appropriate play actions per session. When we introduced the first generalization play set (house), independent appropriate play actions decreased, but then returned to treatment levels when we introduced the error correction procedure. For the final two generalization play sets, the number of independent appropriate play actions per session was at or above treatment levels.

Jude

Jude completed up to 44 appropriate play actions during baseline with the pirate plat set, and did not complete any appropriate play actions with the other two play sets. At the beginning of the multiple exemplar training plus discrimination training, he completed up to seven independent appropriate play actions. Following the series of procedural modifications (around session 64), we observed a steady increase in independent appropriate play actions, and around session 96 Jude's independent appropriate play actions increased sharply up to 53 independent actions at session 101. During the last 20 sessions of the multiple exemplar alone phase, Jude completed between 10 and 53 independent appropriate play actions. When we tested for generalization of varied play behavior to the first set, Jude completed between 21 and 36 independent appropriate play actions. With the second and third sets, he completed between 31 and 42 independent appropriate play actions.

CHAPTER V

DISCUSSION

The purpose of this dissertation was to evaluate the effects of multiple exemplar training plus discrimination training on (a) increasing varied play behaviors of children with autism, (b) producing discriminated responding of varied and repetitious play behavior with trained play sets, and (b) promoting generalization of varied play behavior to untrained play sets. We evaluated the effects of lag schedules of reinforcement plus error correction on increasing varied play behavior with all three participants, and found that, following some procedural modifications, these procedures were effective at increasing varied appropriate play behaviors for all three. We evaluated the effects of multiple exemplar training plus discrimination training with two participants, and did not observe discriminated responding of varied play behavior. For all three participants, we evaluated the effects of multiple exemplar training alone on promoting generalization of varied play behavior to untrained play sets. In general, we found that multiple exemplar training alone was effective at promoting generalization of varied play behavior to at least some untrained play sets for each participant. More specifically, multiple exemplar training alone after conducting discrimination training was effective at promoting generalization of varied play behavior to at least some untrained play sets for two participants. Multiple exemplar training alone (without a prior history of discrimination training) was effective at promoting generalization of varied play behavior to all untrained play sets for one participant.

Effects of Lag Schedules of Reinforcement plus Error Correction

For all three participants, lag schedules of reinforcement plus error correction implemented across multiple play sets (i.e., multiple exemplar training alone) was effective at producing increases in varied play behavior across all three play sets. This is supported by the increases observed in the number of independent appropriate play actions meeting the Lag 1 (Leonard and Arthur) and Lag 2 (Jude) schedules and by the increases in the number of different independent appropriate play actions completed each session. However, we did not observe increases in varied play behavior until we made multiple procedural modifications. We did not observe increases in Leonard's independently varied play behavior until after implementing multiple changes to his reinforcement and prompting procedures, and discontinuing the discrimination training component of the study. When we discontinued the discrimination training component for Leonard, we saw an increase in varied play behavior with one play set (farm), but varied play remained low in the other two play sets. When we introduced the "prompt to action" modification, we saw increases in all three play sets. For Arthur, we made one procedural modification (increased wait time to 20 s), but we did not observe increases in his varied play behavior until we discontinued the discrimination training component. For Jude, who was not exposed to the discrimination training procedures, the multiple exemplar training alone did not appear to initially be effective at increasing his varied play behavior. We did not observe increases in his varied play until after we implemented a series of modifications to his reinforcement procedures.

Given the amount of changes made throughout the study, it is difficult to say

exactly what set of procedures is responsible for the increases in the participants' varied play behavior. It is possible that if we had included all of these modifications in the procedures from the beginning, we might have observed higher levels of varied play behavior sooner, and for Leonard and Arthur we may even have observed discrimination between varied and repetitious play behavior. Therefore, we can only conclude that the multiple exemplar training consisting of lag schedules of reinforcement with error correction was effective at increases the participants' varied play behavior following these specific series of changes. In general, this weakens the conclusions we are able to draw regarding the effectiveness of the lag schedules of reinforcement with error correction at increasing varied play behavior. However, the core components of the procedures remained the same across all three participants. That is, for all three, reinforcement was delivered on a lag schedule, error correction was implemented if the participants' independent behavior did not meet the schedule of reinforcement, and these procedures were implemented across three play sets. We made changes to the way we implemented each of these components, rather than changing the components themselves. For example, we made changes to the stimuli delivered as reinforcers, while maintaining the lag schedules of reinforcement. We made changes to the types of prompts used, while maintaining prompts and error correction when the participants' behavior did not contact the schedule of reinforcement. Thus, to an extent, we are able to conclude that multiple exemplar training consisting of lag schedules of reinforcement and error correction was effective at increasing varied play behavior.

It is worth noting here that there were differences in the way the participants

responded to the lag schedules during the multiple exemplar training alone phase (irrespective of the procedural modifications). For Leonard and Arthur, as the number of independent appropriate responses meeting the Lag 1 schedule increased, so did the number of different responses completed each session. For these two participants, these two dependent measures tended to correspond and increase together. Jude's pattern of responding was different in that, while the number of independent play actions increased drastically, the number of different responses completed each session did not. That is, while Leonard and Arthur varied their behavior across a large number of different play actions each session, Jude varied his behavior across only a few different play actions each session. These data suggest that Jude may have been engaging in higher order stereotypy. Inspection of the raw data confirmed this. During most sessions, Jude rotated across the same three play actions in such a way as to contact the Lag 2 contingency. Within some of these sessions, he rotated across one set of three actions, and then switched to a different set of three actions (thus completing up to six or more different actions per session). Higher order stereotypy is a commonly noted limitation to using lag schedules (Lee, Sturmey, & Fields, 2007). One recommendation for possibly addressing higher order stereotypy is to increase the lag schedule (Contreras & Betz, 2016), thus requiring the participant to complete a larger number of different responses in order to contact reinforcement. However, addressing higher order stereotypy was outside the scope of this study. Future researchers might consider conducting a study specifically aimed at testing methods for addressing higher order stereotypy.

An interesting aspect of the data worth noting is the pattern observed in the

overall number of independent appropriate play actions completed each session by each participant. In general, we found that on sessions where the participants completed few responses that met the lag schedule or were different, they completed few independent appropriate play actions at all. Thus, low levels of variability were not necessarily indicative of invariant play behavior, but rather of low levels of appropriate play behavior in general. As the number of play actions meeting the lag schedules increased, so did the total number of independent appropriate play actions. This suggests that, while we were promoting varied play behaviors, we were doing so while promoting increases in independent appropriate play. These data replicate the Harris (2016) study and highlight the importance of the types of behaviors that are being targeted for increases in variability. As discussed in the literature review, increasing the variability of any play behavior may result in maladaptive play behaviors. Thus, it is important to define parameters of the behaviors that will result in reinforcement such that appropriate and contextual behaviors are being selected for. This will result in increases in variability of play behaviors that can then be used to further promote language and other social learning opportunities.

Another interesting aspect of the data are the cumulative number of novel independent appropriate play actions that each participant completed throughout the study. During the multiple exemplar plus discrimination training phase of the study, Leonard and Arthur completed few novel play actions (e.g., a total of 17). When we discontinued the discrimination training component of the study, such that the participant's appropriate play behavior was exposed to the Lag 1 schedule of

reinforcement only (rather than alternating between the Lag and Rep schedules), we saw a large increase in the completion of novel appropriate play actions. Leonard completed a total of 50 novel actions, and Arthur completed a total of 53 novel play actions. Even Jude, who was engaging in higher order stereotypy by alternating his responding across a small number of different play actions each session, completed a total of 49 novel play actions across the study. These data provide additional support to our conclusion that the lag schedules plus error correction were effective at increasing varied play behavior. More specifically, and perhaps more interesting, these data suggest that not only did the lag schedules plus error correction increase varied play behavior within sessions, but increased varied play behavior across sessions. That is, these procedures resulted in increases in varied play behavior on a larger scale than just within sessions.

Although we observed increases in varied appropriate play, we are unable to draw conclusions regarding the social validity or appropriateness of the levels of varied play we observed with these participants. That is, we do not have comparative data on the levels of appropriate play or varied play with typically developing children. As such, we are unable to speak to whether the increases in varied appropriate play produced in this study are within the range of what we would expect from same-aged typically developing peers. Our results suggest that the procedures used in this study are effective at increasing varied appropriate play. It may be a worthwhile endeavor to investigate how much varied play is “normal”, and then design our procedures to produce levels of varied play within that range. This would increase the social validity of these procedures.

Effects of Multiple Exemplar Training plus Discrimination Training

For Leonard and Arthur, multiple exemplar training plus discrimination training was not effective at producing discriminated responding of varied and repetitious play behaviors. This is, we did not observe clear and large increases in varied responding during the Vary condition over varied responding in the Rep condition for either participant. Upon close inspection of the data, it looks like Leonard and Arthur were beginning to acquire the discrimination, as varied responding was slightly elevated in the Vary component relative to the Rep component. However, variability within the Vary component was generally low throughout this phase of the study for both participants. Also, we observed increases in varied responding within the Rep component for both participants towards the end of this phase. We made multiple procedural modifications for both Leonard and Arthur during this phase, but we still did not observe increases in varied play actions in the Vary component. Given these data, we concluded that the discrimination training was not sufficient to promote a clear discrimination in responding to the two discriminative stimuli. Also, we concluded that the discrimination training may have been preventing Leonard and Arthur from acquiring the skill of varying their behavior with these play sets. This is supported, in part, by the increases in varied responding that we observed when we discontinued the discrimination training portion of the study.

There are several possible reasons why the multiple exemplar training plus discrimination training did not result in discriminated variability. The procedures themselves may not have been effective at producing the discrimination. The

discriminative stimuli (the colored cards and bracelets) may not have been salient enough to signal the reinforcement contingencies in effect for each component. It is also possible that the reinforcement contingencies were not clear enough for the participants' behavior to come under control of the two different schedules. One reason for this may have been our repetition contingency. We allowed for the participants to self-select the response that would result in reinforcement for each Rep session. This procedure allowed the responses reinforced within each Rep sessions to vary across sessions. It may have been clearer to the participants if we had selected one response that would result in reinforcement within all Rep sessions across the entire study.

As discussed in the introduction, Miller (2014) found that discrimination training resulted in discriminated variability of block building with three children with autism, but this discrimination was not strong. One of the purposes of this study was to address the limitations of the procedures used in the Miller study to promote a stronger discrimination. Specifically, we used more salient stimuli as the discriminative stimuli in this study (i.e., visual vs. auditory), and we attempted to make the reinforcement contingencies stricter and thus more clear (i.e., Lag 1 or 2 vs. Lag 3, and self-selected Rep 1 vs. Rep 3). It appears that these modifications were not effective at producing a strong discrimination between variability and repetition. However, there are other factors that may have contributed to the lack of discrimination. It is also possible that conducting the discrimination training across multiple play sets interfered with the acquisition of the discrimination. Future researchers may wish to conduct discrimination training across just one play set at a time in order to evaluate the effects of our discrimination training

procedures alone on (a) producing discriminated variability of play behavior and (b) promoting generalization of the discrimination to untrained play sets. This could perhaps be done in a sequential modification format, such that the discrimination is still trained across multiple play sets. That is, once the discrimination is acquired, the researchers could test for generalization of the discrimination to one play set. If the discrimination is not observed, then it could be trained with this play set, and then tested to another, and so on until generalization of the discrimination is observed.

Generalization of Varied Play Behavior

For Leonard and Arthur, it appears that multiple exemplar training alone following multiple exemplar training plus discrimination training was effective at promoting varied play behavior to two untrained play sets each. This is supported by the moderate to high levels of independent play behaviors that were different and met the Lag 1 schedule with the untrained play sets in the absence of the error correction procedures. Additional support is seen in the steep slope of the cumulative number of novel play actions completed with each generalization play set for Leonard and Arthur. However, the pattern of generalization to the untrained play sets was different for each participant. Leonard engaged in varied play with the first two generalization play sets (pirate ship and Lion King), but not with the third (Paw Patrol) until the error correction procedure was introduced. These data suggest that Leonard's varied play behavior generalized to two sets and not to the third. It is possible, however, that Leonard's behavior with the Paw Patrol play set was idiosyncratic and not generally representative

of his varied play behavior. During the generalization sessions with the Paw Patrol play set, Leonard spent much of his time opening and closing the trap door and spinning the two spin doors. It is possible that these features of the play set allowed Leonard to engage in stereotypic behavior that was more reinforcing than our programmed consequences. It should be noted, though, that the pirate ship play set also had a trap door and the Lion King play set also had a spinning door. Thus, these features were not idiosyncratic to just the Paw Patrol play set. Regardless, he did not engage in varied appropriate play with the Paw Patrol play set in the absence of error correction.

We also observed generalization of Arthur's varied play behavior to two play sets, but the pattern of his responding was different from Leonard's. Arthur initially began to engage in varied play behavior with the first generalization play set (house), but his behavior quickly decreased in the absence of the error correction procedure. After introducing the error correction procedure with the house play set, Arthur's varied play behavior increased to levels similar to that observed in the multiple exemplar training alone. Following error correction with just one of the generalization play sets, Arthur engaged in varied play behavior with the remaining two generalization play sets (farm and Paw Patrol) in the absence of the error correction procedure. This suggests that multiple exemplar training alone, following multiple exemplar training plus discrimination training, was not effective at promoting generalization to untrained play sets with Arthur until the error correction procedure was put in place with the first generalization play set. This series of procedures resembles sequential modification, in that we tested for generalization in the absence of the error correction procedure and then

applied the procedure after observing that generalization of varied behavior did not occur. Thus, for Arthur, it may be more accurate to conclude that the multiple exemplar training alone was effective at promoting generalization of Arthur's varied play behavior when followed by brief sequential modification.

Another way to conceptualize the pattern of responding observed with Arthur is that his varied play behavior initially generalized to the house play set (generalization play set 1) but did not maintain. One potential reason that Arthur's behavior did not maintain is that he may have become satiated (or bored, to speak loosely) with the house play set. During the training phase, we ran between four and six sessions a day, rotating across the three play sets such that he would play with each one up to two times. During the generalization phase of the study, we continued to run between four and six sessions per day, but they were all with the same play set. It is possible that repeated exposure to the same play set resulted in the steady decrease in responding observed with this play set.

It is important to remember that the Arthur was exposed to the Paw Patrol play set for one session under both the Vary and Rep conditions. It is possible that his history with this play set contributed to his varied play behavior with this play set under the generalization condition. However, the Paw Patrol play set was implemented in the treatment procedures early in the study, under the multiple exemplar plus discrimination training phase, where Arthur was not yet engaging in high levels of varied appropriate play behaviors. Additionally, this play set was exposed to the Vary reinforcement contingency during only one session. Thus, it is unlikely that his prior experience with

the Paw Patrol play set was the primary contributor to the high levels varied play behavior with this set.

For Jude, it appears that multiple exemplar training alone, without the history of discrimination training, was effective at promoting generalization of varied play behavior to all three untrained play sets. This is supported by the high levels of independent appropriate play behaviors that met the Lag 2 schedule in the absence of the error correction procedures for all three generalization play sets. It is interesting to note that we did not observe as steep of a slope in the cumulative number of novel actions Jude completed with the generalization play sets. We did observe what could be termed higher order stereotypy during these sessions, which may have prevented him from emitting high levels of novel play actions within the relatively short amount of time spent with these play sets. It is also interesting to note that Jude was the only participant who was not exposed to the multiple exemplar training plus discrimination training phase and is also the only participant whose varied play behavior generalized to all three untrained play sets. For both Leonard and Arthur, we are unable to draw conclusions regarding the effects of just the multiple exemplar training alone on promoting generalization of varied play behavior because this phase followed the multiple exemplar training plus discrimination training phase. It is possible that this history of discrimination training, although unsuccessful at producing a clear discrimination, contributed to the generalization of these participants' varied play behaviors. However, it is also possible that the history of discrimination training impeded generalization to the untrained play sets. This is an empirical question. Future researchers might consider investigating the

effects of multiple exemplar training alone on promoting generalization of varied play behavior.

Considering the different patterns of generalization of varied play behavior across the three participants, it might be useful to think about generalization as a continuum rather than something that is dichotomous and either does or does not happen. On one end of the continuum would be "weak" generalization, where high amounts of varied play behavior did not immediately occur, but rather were acquired over the course of several sessions after coming in contact with the reinforcement contingencies and error correction (generalization set 1 with Arthur and 3 with Leonard). This could be considered generalization even though the error correction procedure was in place because the participants acquired the skill of varying their play behavior much more rapidly with these play sets than they had during the initial training. On the other end of the continuum would be "strong" generalization, where high levels of varied play behavior were observed immediately with the untrained play sets (all generalization sets with Jude and sets 2 and 3 with Arthur). In the middle of the continuum would be "moderate" generalization, where increases in varied play were observed to increase following a few sessions of contacting the reinforcement contingency in the absence of the error correction procedure (generalization sets 1 and 2 with Leonard). Given this conceptualization of generalization, we could conclude that the varied play behavior of all three participants generalized in varying degrees to all of the untrained play sets.

Limitations and Directions for Future Research

There are limitations to this dissertation that should be discussed. As discussed

above, one limitation is that our procedures failed to promote discrimination between varied and repetitive play behavior. Future researchers may wish to further investigate methods for establishing discriminative control over varied play behavior, as well as investigate how a discriminative stimulus for variability could be used to promote generalization of variability to untrained play situations.

Another limitation to this dissertation is the number of procedural modifications that were made across all three participants. We made four modifications to the procedures for Leonard, one for Arthur, and five for Jude. Because of these modifications, the procedures were slightly different for all three participants, thus limiting our ability to draw firm conclusion regarding the effectiveness of any one set of procedures. However, as discussed earlier, the core components of the procedures remained the same across all three participants, although the exact implementation of each component was modified across participants. As the main question of this study was whether we could promote generalization of play variability to untrained play sets, it was vital that we achieve increases in play variability. The modifications we made were necessary towards this end. The procedural modifications we made fell into two categories; (a) changes to reinforcement and (b) changes to prompting procedures. For Leonard and Jude, we made changes to the reinforcement procedures (e.g., added in social attention). For Leonard and Arthur, we made changes to the prompting procedures (e.g., "prompt to action"). In order to more clearly test the effects of the lag schedules themselves on increasing play behavior, it might be useful for future researchers to conduct reinforcer assessments prior to implementing a study, in order to identify stimuli

(or combinations of stimuli) that are more likely to function as reinforcers. It might also be useful to conduct assessments of prompting procedures to determine which might be the most effective at prompting appropriate play actions while minimizing prompt dependency. This would allow the researchers to include effective reinforcement and prompting consistently throughout the entirety of a study, which would allow them to draw firm conclusions regarding the procedures used.

Another aspect of the procedures that might warrant further investigation is the possibility that the error correction procedure used functioned as an aversive stimulus that resulted in a negative reinforcement effect. That is, it is possible that the participants not only responded to the programmed positive reinforcement (i.e., edibles and social interaction delivered on a lag schedule), but were also responding to an unintentional negative reinforcement contingency associated with the error correction procedure. It is possible that the error correction procedure we had in place was aversive, and that the participants' high levels of varied play during the training were a product of escape from the prompting procedure (negative reinforcement) rather than (or in addition to) access to the edibles. This possible effect was observed with Jude, who frequently resisted our physical prompts throughout the study. Around session 96, where we observed a sharp increase in the number of play actions meeting the Lag 2 schedule, we also noticed a decrease in the amount of prompts delivered (which we confirmed within the raw data). It is possible that it was the combination of access to the plate of edibles plus escape from physical prompts that resulted in Jude's high levels of varied play behavior.

Another potential limitation to this study is that we did not conduct component

analyses of the play actions completed in relation to the play sets. We also did not analyze independent play actions in relation to prompted play actions. Conducting more thorough analyses of the raw the data might reveal patterns that could be used to further promote acquisition of varied play behavior and generalization of varied play behavior. For example, if there were features of some play sets that seemed to be associated with certain play actions, those features of the play sets could then be incorporated into the generalization play sets to further promote generalization. If there were patterns in the types of actions that were being prompted versus independent play actions, this information could be used to help guide the way the therapists prompt in order to further promote novel responding. Future researchers might consider conducting component analyses of varied play behavior. In summary, the results of this dissertation extend the previous research on generalization of varied play behaviors in that it is the first to successfully promote generalization of varied play behavior to untrained play situations. These data suggest that it is possible to promote generalization of response variability in the context of play, and that multiple exemplar training may be an effective strategy for doing so. However, there is still much work to be done and future researchers should continue to investigate methods for promoting generalization of response variability. Researchers should work to refine the methods for promoting generalization of variability of play behavior, as well as explore methods for promoting generalization among other operants (e.g., verbal behavior). Also, applied researchers should investigate ways to bring methods for promoting generalization of varied behavior into the everyday treatment of individuals with autism.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5*th ed.). Washington, DC: Author.

Baron-Cohen, S. (1987). Autism and symbolic play. *British Journal of Developmental Psychology, 5,* 139-148.

Barton, E. E., & Wolery, M. (2008). Teaching pretend play to children with disabilities: A review of the literature. *Topics in Early Childhood Special Education, 28(2),* 109-125.

Baruni, R. R., Rapp, J, T., Lipe, S. T., & Novotny, M. A. (2014). Using lag schedules to increase toy play variability for children with intellectual disabilities. *Behavioral Interventions, 29,* 21-35.

Betz, A. M., Higbee, T. S., Kelley, K. N., Sellers, T. P., & Pollard, J. S. (2011). Increasing response variability of mand frames with script training and extinction. *Journal of Applied Behavior Analysis, 44(2),* 357-362.

Brodhead, M. T., Higbee, T. S., Gerencser. K. R., & Akers, J. S. (2016). The use of a discrimination-training procedure to teach mand variability to children with autism. *Journal of Applied Behavior Analysis, 49(1),* 34-48.

Buysse, V., Wesley, P, Keyes, L., & Bailey, D. B. (1996). Assessing the comfort zone of child care teachers in serving young children with disabilities. *Journal of Early Intervention, 20(3),* 189-203.

Cammilleri, A. P., & Hanley, G. P. (2005). Use of a lag differential reinforcement contingency to increase varied selections of classroom activities. *Journal of Applied Behavior Analysis, 38(1),* 111-115.

Charman, T., & Baron-Cohen, S. (1997). Brief report: Prompted pretend play in autism. *Journal of Autism and Developmental Disorders, 27*(3), 325-332.

Contreras, B. C., & Betz, A. M. (2016). Using lag schedules to strengthen the intraverbal repertoires of children with autism. *Journal of Applied Behavior Analysis, 49(1),* 3-16.

Esch, J. W., Esch, B. E., & Love, J. R. (2009). Increasing vocal variability in children with autism using a lag schedule of reinforcement. *The Analysis of Verbal Behavior, 25,* 73-78.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25(2),* 491-498.

Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis, 27(4),* 739-760.

Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and the emergence of new forms in children's blockbuilding. *Journal of Applied Behavior Analysis, 6(2),* 209-217.

Harding, J. W., Wacker, D. P., Berg, W. K., Gary, R., & Lee, J. F. (2004). Promoting response variability and stimulus generalization in martial arts training. *Journal of Applied Behavior Analysis, 37(2),* 185-195.

Harris, M. K. E. (2016). *An analysis of variability of play behavior with preschool children with autism* (Doctoral dissertation). Utah State University, Logan, UT. Retrieved from http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=6009&context=etd

Heldt, J., & Schlinger, H. D. (2012). Increased variability in tacting under a Lag 3 schedule of reinforcement. *The Analysis of Verbal Behavior, 28,* 131-136.

Holman, J., Goetz, E. M., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In B. Etzel, J. LeBlanc, & D. Baer (Eds.), *New developments in behavioral research.* (pp. 441-471). Hillsdale, NJ: Erlbaum.

Jarrold, C., Boucher, J. & Smith, P. K. (1996). Generativity deficits in pretend play in autism. *British Journal of Developmental Psychology, 14,* 275-300.

Koehler-Platten, K., Grow, L. L., Schulze, K. A., & Bertone, T. (2013). Using a lag reinforcement schedule to increase phonemic variability in children with autism spectrum disorders. *The Analysis of Verbal Behavior, 29,* 71-83.

Lalli, J. S., Zanolli, K., & Whon, T. (1994). Using extinction to promote response variability in toy play. *Journal of Applied Behavior Analysis, 27(4),* 753-736.

Lee, R., McComas, J. J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *Journal of Applied Behavior Analysis, 35(4),* 391-402.

Lee, R., & Sturmey, P. (2006). The effects of lag schedules and preferred materials on variable responding in students with autism. *Journal of Autism and Developmental Disorders, 36(3),* 421-428.

Lee, R., & Sturmey, P. (2014). The effects of script-fading and a Lag-1 schedule on varied social responding in children with autism. *Research in Autism Spectrum Disorders, 8,* 440-448.

Lee, R., Sturmey, P., & Fields, L. (2007). Schedule-induced and operant mechanisms that influence response variability: A review and implications for future investigations. *The Psychological Record, 57,* 429-455.

Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior, 52(2),* 155-166.

McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders, 32(5),* 351-372.

McCune, L. (1995). A normative study of representational play at the transition to language. *Developmental Psychology, 31(2),* 198-206.

Miller, N. (2014). *Stimulus control and generalization of operant variability in the block play of children with autism* (Doctoral dissertation). The Ohio State University, Columbus, OH

Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis, 33(2),* 151-165.

Murray, C., & Healy, O. (2012). Increasing response variability in children with autism spectrum disorder using lag schedules of reinforcement. *Research in Autism Spectrum Disorders, 7,* 1481-1488.

Napolitano, D. A., Smith, T., Zarcone, J. R., Goodkin, K., & McAdam, D. B. (2000). Increasing response diversity in children with autism. *Journal of Applied Behavior Analysis, 43(2),* 265-271.

Neuringer, A. (2004). Reinforced variability in animals and people. *American Psychologist, 59(9),* 891-906.

Newman, B., Reinecke, D. R., & Meinberg, D. L. (2000). Self-management of varied responding in three students with autism. *Behavioral Interventions, 15,* 145-151.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology, 11(3),* 429-452.

Parsonson, B. S., & Baer, D. M. (1978). Training generalized improvisation of tools by preschool children. *Journal of Applied Behavior Analysis, 11(3),* 363-380.

Pryor, K., Haag, R., & O'Reilly, J. (1969). The creative porpoise: training for novel behavior. *Journal of the Experimental Analysis of Behavior, 12,* 653-661.

Rodriguez, N. M., & Thompson, R. H. (2015). Behavioral variability and autism spectrum disorders. *Journal of Applied Behavior Analysis, 48(1),* 167-187.

Sellers, T. P., Kelley, K., Higbee, T. S., & Wolfe, K. (2016). Effects of simultaneous script training on use of varied mand frames by preschoolers with autism. *The Analysis of Verbal Behavior, 32,* 15-26.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10(2),* 349-367.

Susa, C., & Schlinger, H. D. (2012). Using a lag schedule to increase variability of verbal responding in an individual with autism. *The Analysis of Verbal Behavior, 28,* 125-130.

Williams, E., Reddy, V., & Costall, A. (2001). Taking a closer look at functional play in children with autism. *Journal of Autism and Developmental Disorders, 31*(1), 67-77.

Wiskow, K. M., & Donaldson, J. M. (2016). Evaluation of a lag schedule of reinforcement in a group contingency to promote varied naming of categories items with children. *Journal of Applied Behavior Analysis, 49(3),* 1-13.

Wolfe, K., Slocum, T. A., & Kunnavatana, S. S. (2014). Promoting behavioral variability in individuals with autism spectrum disorders: A literature review. *Focus on Autism and Other Developmental Disabilities, 29(3),* 180-190.

CURRICULUM VITAE

BETHANY P. CONTRERAS

EDUCATION

Utah State University
Ph.D. in Disability Disciplines, Applied Behavior Analysis Specialization
May 2017
Advisor: Thomas Higbee, Ph.D., BCBA-D
Dissertation: Evaluation of Multiple Exemplar Training plus Discrimination Training to Promote Generalization of Response Variability

Florida Institute of Technology
M.S. in Applied Behavior Analysis and Organizational Behavior Management
August 2012
Advisor: Alison M. Betz, Ph.D., BCBA-D
Thesis: Using Lag Schedules to Strengthen the Intraverbal Repertoires of Children with Autism

Arizona State University
B.S., Biology (general)
August 2008

PROFESSIONAL LICENSES AND CERTIFICATIONS

Utah Licensed Behavior Analyst (LBA)
September 2015-present
License number: 9536978-2506

Board Certified Behavior Analyst (BCBA)
September 2012-present
Certification number: 1-12-12534

AWARDS AND SCHOLARSHIPS

- Florida Tech Honors Convocation Outstanding Graduate Student in Applied Behavior Analysis, 2012
- Florida Tech Graduate Student Assistant (2), 2011-2012
- Florida Tech Graduate Scholarship, 2010-2011
- Arizona State University, Magna Cum Lade
- Arizona State Dean's List for Academic Achievements, GPA of 3.5 or higher, 2004-2008

TEACHING EXPERIENCE

Utah State University

- SPED 5010: Applied Behavior Analysis 1: Principles, Assessment, and Analysis
  Primary Instructor for the undergraduate distance course
  Responsibilities: prepared course content, delivered lectures, held office hours, graded assignments
  Fall 2015

- SPED 5010: Applied Behavior Analysis 1: Principles, Assessment, and Analysis
  Teacher's Assistant for the undergraduate on-campus course
  Responsibilities: graded assignments, assisted with course prep, held office hours, delivered multiple lectures
  Fall 2013, Fall 2014, Fall 2015

- PSY 3210: Abnormal Psychology
  Undergraduate on-campus course
  Guest lecture on applied behavior analysis and developmental disabilities
  Fall 2015

- SPED 5780: Foundations in Special Education and Legal Issues
  Undergraduate distance course
  Guest lecture on Strategies for Supporting Independence in Individuals with Autism
  Fall 2014, Fall 2015

- Communication Disorders Master's Lecture Series
  Master's on-campus course
  Guest lectures on Understanding Behavior: General Strategies to Support Success and Understanding and Managing Challenging Behavior
  Fall 2013, Fall 2014, Fall 2016

- SPED 6710: Concepts and Principles of Behavior Analysis in Education
  Teacher's Assistant for Master's distance course
  Responsibilities: graded assignments, assisted with course prep, held office hours, delivered multiple lectures
  Spring 2014

Florida Institute of Technology

- BEHP 5002: Behavior Change Procedures and Ethical Considerations, Florida Institute of Technology
  Co-instructor for Master's distance course
  Responsibilities: design and prepare course content, deliver lectures, provide support to students
  2013

- PSY 3421: Learning and Motivation
  Teacher's Assistant for undergraduate on-campus section
  Responsibilities: held office hours, created testing materials, graded assignments, delivered multiple lectures
  Fall 2011, Spring 2012

Arizona State University

- Bio 201/202: Human Anatomy and Physiology 1 and 2
  Teacher's assistant for undergraduate on-campus lab sections
  Responsibilities: assist students during lab, assist in creating tests and study guides, presented multiple lectures
  Spring 2006, Fall 2007

PUBLISHED MANUSCRIPTS

- **Contreras, B. P.**, & Betz, A. M. (2016). Using lag schedules to strengthen the intraverbal repertoires of children with autism. *Journal of Applied Behavior Analysis, 49(1),* 1-14.
- Hoffman, A. M., **Contreras, B. P.,** & Clay, C. J. (2016). Acceptance and commitment therapy for individuals with disabilities: A behavior analytic strategy for addressing private events in challenging behavior. *Behavior Analysis in Practice, 9(1),* 14-24.
- Loughrey, T., **Contreras, B. P.,** Blank, P., Rudy, N., Majdalany, L., Sinn, S., Coxon, G., Harvey, A., & McGreevy, P. (2014). Training caregivers as interventionists to teach mands and as trainers to teach family members mand training procedures. *The Analysis of Verbal Behavior, 30(2),* 128-140.
- Gerencser, K. R., Higbee, T. S., Akers, J. A., & **Contreras, B. P.** (in press). Evaluation of interactive computerized training to teach parents to implement photographic activity schedules with children with autism spectrum disorder. *Journal of Applied Behavior Analysis.*

MANUSCRIPTS IN PREPARATION

- **Contreras, B. P.,** Higbee, T. S., Hess, B., & Gunn, S. (under review). Using an activity schedule to promote independent eating: A data-based case study.
- **Contreras, B. P.,** Slocum, T. A., Higbee, T. S. (in preparation). Effects of Errorless Teaching Procedures on Learning Outcomes for Individuals with Autism: A Systematic Review.
- **Contreras, B. P.**, Miller, S. A., & Betz, A. M. (in preparation). Signaled delay to reinforcement in a multiple schedule training to reduce excessive manding.
- Gerencser, K. R., Higbee, T. S., **Contreras, B. P**. & Pellegrino, A. J. (in preparation). Evaluation of an interactive computerized training to teach paraprofessionals to implement errorless discrete trial instruction.

PROFESSIONAL PRESENTATIONS

- **Contreras, B. P.,** Reinert, K., & Higbee, T. S. (2016). Advanced reinforcement systems. Break-out presentation at the annual conference for the Utah Multi-tiered Systems of Support Connections Conference, Layton, UT.

- Gunn, S. V., & **Contreras, B. P.** (2016). Multi-tiered approach to increasing self-help skills in early childhood settings. Break-out presentation at the annual conference for the Utah Multi-tiered Systems of Support Connections Conference, Layton, UT.

- **Contreras, B. P.,** Slocum, T. A., & Higbee, T. S. (2016). Effects of Errorless Teaching Procedures on Learning Outcomes for Individuals with Autism: A Systematic Review. Presentation in J. Ninci (Chair), Instructional Practices for Individuals with Autism Spectrum Disorder: Syntheses of the Literature. Symposium Presented at the annual conference for the Association of Behavior Analysis International, Chicago, IL.

- **Contreras, B. P.,** Higbee, T. S., Hess, B., & Gunn, S. V. (2016). Using activity schedules to promote independent eating in children with autism: a pilot study. Poster presented at the Annual Conference for the Utah Association for Behavior Analysis, Salt Lake City, UT.

- **Contreras, B. P.,** Gerencser, K. R., Higbee, T. S. (2014). Systematic compliance training for children with autism. Poster presented at the Annual Conference for the Utah Association for Behavior Analysis, Layton, UT.

- **Contreras, B. P.,** Betz, A. M. (2012) Effects of lag schedules on variability of responding to Intraverbal questions with children with autism. Presentation in A. Betz (Chair), The Effects of Behavioral Interventions to Treat Behaviors Commonly Associated with Autism. Symposium Presented at the Annual Florida Tech School of Psychology Research Colloquium, Melbourne FL.

- **Contreras, B. P.,** Betz, A. M. (2012) Effects of lag schedules on variability of responding to Intraverbal questions with children with autism. Presentation in D. A. Wilder, (Chair), Skill Acquisition in Behavior Analysis: Teaching Safety, Compliance, and Language to Children. Symposium presented at the Annual Conference for the Association for Applied Behavior Analysts, Tampa, FL.

- Miller, S. A., **Contreras, B. P.**, Betz, A. M. (2012). Evaluating transfer of stimulus control during reinforcement schedule thinning. Presentation in K. C. Luczynski (Chair), Enhancing Function-Based Interventions for Problem Behavior Using Multiple Schedules. Symposium presented at the Annual Conference for Association for Behavior Analysis International, Seattle, WA.

- **Contreras, B. P.,** Betz, A. M. (2012) Effects of lag schedules on variability of responding to Intraverbal questions with children with autism. Presentation in R. G. Miltenberger, (Chair), Recent Research on Teaching Verbal Behavior to Children with Autism. Symposium presented at the Annual Conference for the Florida Association for Behavior Analysis, Jacksonville, FL.

- Betz, A. M. **Contreras, B. P.**, Miller, S. A. (2012). Assessing variations of multiple schedule arrangements to teach children with autism to tolerate delays to reinforcement. Presentation in A. M. Betz (Chair) Teaching Socially Significant Behaviors: From Self Control to Social Interactions. Symposium presented at the Annual Conference for the Florida Association for Behavior Analysis, Jacksonville, FL.

- Loughrey, T., **Contreras, B. P.,** Blank, P., Rudy, N., Majdalany, L., Sinn, S., Coxon, G., Harvey, A., & McGreevy, P (2012). Preparing caregivers as interventionists and trainers: Teaching Verbal behavior to children with developmental disabilities. Presentation in J. Oliver (Chair) Interventions from the Edge: Education and Treatment in Home, School, and Clinical Settings. Symposium presented at the Annual Conference for the Florida Association for Behavior Analysis.

- **Contreras, B. P.**, Miller, S. A. (2012). Decreasing steps per task through ergonomic design at a locally owned coffee shop. Poster presented at the Annual Conference for the Association for Behavior Analysis International, Seattle, WA.

- **Contreras, B. P.**, & Betz, A. M. (2012) Effects of lag schedules on variability of responding to intraverbal questions with children with autism. Poster presented at the Annual Conference for the Association for Behavior Analysis International, Seattle, WA.

- **Contreras, B. P.**, Miller, S. A. (2011). Decreasing steps per task through ergonomic design at a locally owned coffee shop. Poster presented at the Annual Conference for the Florida Association for Behavior Analysis, Daytona, FL

- Miller, S. A., **Contreras, B. P.,** Betz, A. M. (2011). Signaled delay to reinforcement in a multiple schedule training to reduce excessive manding. Poster presented at the Annual Conference for the Florida Association for Behavior Analysis, Daytona, FL.

CLINICAL EXPERIENCE

- *Behavior Analysis Consultant to Granite School District, Autism Support Services: Education, Research, and Training (ASSERT)*, August 2016-present
  - Responsibilities: provide supervision and training to special education teachers in Granite School District (Salt Lake City, UT), provide case management support in multiple ABA focus classrooms

- *Behavior Analyst*, *Autism Support Services: Education, Research, and Training (ASSERT),* August 2013 – August 2016
  - Responsibilities: managed programming for multiple students; managed team of instructors; organized performance management for ASSERT staff; presented staff, professional, parent, and community trainings and lectures

- *Behavior Analyst, Scott Center for Autism Treatment,* November 2012-February 2013
  - Responsibilities: managed programming for multiple students, managed team of instructors, supported performance management for Scott Center staff

- *Floor Supervisor, Scott Center for Autism Treatment,* May 2012-Novemeber 2012
  - Responsibilities: provided ongoing performance management and support to floor staff, assisted in development of content for the "Learning to Live with Autism" parent training project

- *Primary Behavior Specialist, Scott Center for Autism Treatment*, August 2011-May 2012
  - Responsibilities: provided direct service to multiple children with autism (including DTT, feeding intervention, and functional analysis), assisted with case management

- *Behavior Specialist, Scott Center for Autism Treatment*, May 2011-August 2011
  - Responsibilities: provided direct service to multiple children with autism

- *Behavior Specialist, Behavior Services of Brevard*, July 2009-May 2011
  - Responsibilities: provided direct service to adults with disabilities, provided support for community involvement

- *Lead Behavior Specialist, Arizona Autism United*, January 2007-May 2008
  - Responsibilities: provided direct service to multiple children with autism, assisted with case management

- *Behavior Specialist, Gala Human Services Cooperative*, October 2005-January 2007
  - Responsibilities: provided direct service to multiple children with autism

PROFESSIONAL MEMBERSHIPS

- Association for Behavior Analysis International (ABAI)
- Utah Association for Applied Behavior Analysis (UTABA)
- California Association for Behavior Analysis (CalABA)