Effects of Conditioning Procedures on Vocalizations of Children with Minimal or Emerging Echoic Repertoire

By

Kelley L. Harrison

Submitted to the graduate degree program in Applied Behavioral Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Chairperson Pamela L. Neidert, Ph.D., BCBA-D

____
Claudia L. Dozier, Ph.D., BCBA-D

____
Jan B. Sheldon, Ph.D., J.D.

____
Thomas L. Zane, Ph.D., BCBA-D

____
Jason C. Travers, Ph.D., BCBA-D

Date Defended: June 22nd 2018

The Dissertation Committee for Kelley L. Harrison certifies that this is the approved version of the following dissertation:

Effects of Conditioning Procedures on Vocalizations of Children with Minimal or Emerging Echoic Repertoire

____
Chairperson Pamela L. Neidert, Ph.D., BCBA-D

Date Approved: June 22nd 2018

Abstract

Researchers have utilized conditioning procedures to establish neutral stimuli as conditioned reinforcers for decades (e.g., Greer, Pistoljevic, Cahill, & Du, 2011; Lovaas et al., 1966; Dorow, 1980). More recently, researchers have used conditioning procedures as a strategy for facilitating language acquisition (e.g., Sundberg, Michael, Partington, & Sundberg, 1996). However, the effectiveness of these procedures is unclear. Therefore, Study 1 compared three different procedures suggested to condition speech sounds as automatic reinforcers: stimulus-stimulus pairing (Esch, Carr, & Grow, 2009; Sundberg, Michael, Partington, & Sundberg, 1996), response-stimulus pairing (Lepper and Petursdottir, 2017), and operant discrimination training (Lepper, Petursdottir, & Esch, 2013). Study 2 evaluated potential barriers to the efficacy of conditioning procedures (i.e., articulation, reinforcer efficacy, attending skills) with Study 1 participants for whom none of the conditioning procedures was effective. Multielement, multiple-baseline, and reversal designs were used to demonstrate experimental control. Seven children with and without developmental disabilities participated. Results of Study 1 were idiosyncratic. At least one conditioning procedure increased vocalizations for three participants. An echoic repertoire developed during the evaluation for one participant. None of the three conditioning procedures increased vocalizations for three participants. In Study 2, at least one potential barrier was identified for the three participants. Overall, results suggest that children's vocal behavior may be differentially sensitive to the procedures evaluated. Further, articulation skills, reinforcer efficacy, and attending skills appear to be barriers that may decrease the effectiveness of these procedures for increasing vocalizations.

Keywords: stimulus-stimulus pairing; response-stimulus pairing; operant discrimination training; conditioned reinforcers; automatic reinforcers; discriminative stimulus; respondent

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Pamela Neidert, without whom I would not be receiving my Ph.D.! Thank you for taking a chance on me six years ago, providing me with so many important opportunities, and guiding me through this crazy journey called graduate school. I am so proud to be a Neidert! Second, I would like to thank my committee members, Drs. Claudia Dozier, Jan Sheldon, Thomas Zane, and Jason Travers, for their support and insightful suggestions regarding this project. In particular, I would like to thank Dr. Claudia Dozier for her constant support and guidance throughout my time at KU, even though I wasn't directly her student. It was under her supervision in Little Steps that this project started, and I can't thank her enough for the several rounds of revisions provided that really got this project moving!

I would also like to thank my graduate student colleagues, past and current, who have become life-long friends. I owe a large thank you to Drs. Brian Greer, Joe Dracobly, Adam Briggs, and Jessica Juanico for providing feedback on every stage of this document, and always providing encouragement when needed. Specifically, I would like to thank Dr. Jessica Juancio, who has been my right-hand gal from day one, and without whom I could not have made it to this point. I also owe a large thank you to Alec Bernstein, Marcella Hangen, and Ashley Romero who have been instrumental in the completion of this project. Not only did these three provide feedback on every stage of this document, but they conducted countless sessions, analyzed data, and created countless graphs. I promise – I will never make you listen to my nonsense sounds again!

Finally, I must thank my family, without whom I could not be here today. My mom and dad taught me that anything was possible, and never once doubted I would be here today – even

when I frequently doubted it myself. My sister, Mary Kate, cousins, Kayla and Laura, my grandma and grandpa, and my Aunt Cathy have also offered constant support – listening to countless stories and sending countless messages of encouragement! And, I must give a special thanks to my Aunt Lynne who not only offered constant support and sent countless messages of encouragement (your cards mean so much to me!), but also taught me the difference between active and passive voice, and how to use those commas! I have been extremely blessed to be surrounded by such a special community and provided such amazing opportunities. Looking back, I would never choose a different path. Rock Chalk!

Table of Contents

Effects of Conditioning Procedures on Vocalizations of Children with Minimal or Emerging Echoic Repertoire

Language delays are common in children with intellectual or developmental disabilities (IDD; Patten et al., 2014; Petursdottir & Lepper, 2015). The degree of language impairment can vary across children, with some children exhibiting a complete absence of vocal language (Kjelgaard & Tager-Flusberg, 2001). Language allows children to interact with and interpret their environment (Hart & Risley, 1995). Additionally, language facilitates most learning opportunities for children (e.g., following instructions, imitation, labeling environmental stimuli, requesting environmental stimuli such that daily needs are met; Sundberg & Michael, 2001). Thus, language delays are problematic.

Language delays may occur for several reasons. First, language delays may occur because the behavior of children with IDD is likely to engage in a lower frequency of vocalizations (Patten et al., 2014), and from a behavioral perspective, language develops in part due to respondent conditioning (Shillingsburg, Hollander, Yosick, Bowen, & Muskat, 2015). If children with IDD engage in a lower frequency of vocalizations, then they are likely to contact fewer conditioning opportunities. Second, languages delays may occur because the behavior if children with IDD is likely to be insensitive to social reinforcement (Petursdottir & Lepper, 2015), and from a behavioral perspective, language also develops in part due to operant processes such as automatic (e.g., hearing oneself speak) or differential reinforcement (e.g., parental attention delivered contingent on child vocalizations; Shillingsburg, Hollander, Yosick, Bowen, & Muskat, 2015). Thus, an insensitivity to operant processes may result in a decrease in rate and variability of a child's vocal behavior. Third, language delays may occur because children with IDD may lack the skills to orient to social stimuli such as speech sounds

(Shillingsburg et al., 2015). Orienting toward relevant stimuli is a pre-requisite skill for joint attention, which is correlated with language development (Toth, Munson, Meltzoff, & Dawson, 2006).

Clinicians and researchers have used numerous strategies to facilitate language acquisition (i.e., strategies to increase vocalizations or communication), with common strategies including echoic training (e.g., Stock, Schulze, & Mirenda, 2008), motor imitation training (e.g., Ross & Greer, 2003), mand training (e.g., Hall & Sundberg, 1987), and differential reinforcement of all vocalizations (e.g., Lovaas et al., 1966). These strategies can be problematic, however, because a therapist cannot physically prompt correct speech responses, which often leads to frequent failures for the child. For example, in echoic training, a therapist provides a vocal model and offers reinforcement when the child's vocal sounds match the therapist's vocal sounds (Stock et al., 2008). Because the therapist cannot physically prompt the child to match his or her vocal sounds, the child may fail several times before accurately matching the therapist's vocal sounds to receive reinforcement (Ward et al., 2007). Lovaas (2003) avoided this problem by suggesting that therapists deliver reinforcement immediately following any vocalization emitted by the child. However, this strategy is also problematic if the child infrequently emits vocalizations such that the opportunity to reinforce vocalizations is rare (Petursdottir & Lepper, 2015).

Other researchers have suggested an alternative strategy to facilitate language acquisition (specifically, increasing the frequency of vocalizations) that involves conditioning speech sounds as automatic reinforcers (e.g., Sundberg, Michael, Partington, & Sundberg, 1996). For individuals with significant language deficits (i.e., individuals without an echoic repertoire or individuals who produce few to no speech sounds), procedures to establish vocalizations as automatic reinforcers may be more appropriate because these procedures do not require prompts to ensure correct responding and do not require frequent or varied vocalizations (Ward et al., 2007). To date, three general strategies to establish vocalizations as automatic reinforcers have been described in the applied behavioral literature. These strategies are stimulus-stimulus pairing (SSP; e.g., Sundberg, et al., 1996), operant discrimination training (ODT; e.g., Lepper, Petursdottir, & Esch, 2013), and response-stimulus pairing (RSP; e.g. Lepper & Petursdottir, 2017). However, there are limited experimental analyses of these conditioning procedures. Of the experimental analyses that do exist, researchers report mixed findings with respect to increasing the frequency with which participants emit targeted vocalizations. Therefore, it is unclear whether these procedures should be recommended for increasing vocalizations, and if so, which procedure should be used and under what conditions. Thus, the purpose of the current study is to compare the effects of these three procedures on the level of novel speech sounds emitted by children with a minimal echoic repertoire.

Typical and Delayed Speech Development

With respect to chronological development, when infants are born, they typically engage in what researchers have referred to as "prespeech." That is, they engage in nonfunctional speech that most likely develops as part of the maturation process. This includes crying, cooing, and some babbling (Schlinger, 1995). Prespeech begins at birth and develops into the infant's first language-like sounds in the first two months. First, quasi-vowel sounds develop (e.g., ooh, aah). Then, quasi-vowel sounds with hard consonants begin to develop (e.g., goo). By 5 to 6 months of age, most infants engage in frequent babbling. By 6 to 10 months, canonical babbling (i.e., consonant and vowel sequences) begins, but sounds are not functional. That is, the infants do not use canonical babbling as a way to communicate their wants and needs. However, most caregivers begin to treat this babbling as functional because it is often the first time speech sounds begin to closely resemble the speech sounds of caregivers. Thus, through differential reinforcement, babbling is shaped into functional communication (i.e., infants begin to use babbling as a way to communicate their wants and needs; Schlinger, 1995). Typically, infants begin uttering their first words by year 1 and begin using short sentences by year 2 (Hoff, 2005).

In contrast, children with IDD often experience language delays that become apparent during infancy. Patten et al. (2014) examined vocalizations via video of 37 infants at 9-12 months and at 15-18 months for canonical babbling and rate of vocalizations. Participants included infants from a larger study based on the availability of video recordings, and were not determined to be at risk for ASD before inclusion in the study. Of the 37 infants, 23 were later diagnosed with autism spectrum disorder (ASD). Those who were diagnosed with ASD were less likely to be classified as being in the canonical babbling stage, engaged in less canonical babbling as compared to the typically developing infants, and produced fewer vocalizations as compared to the typically developing infants. If speech develops as a consequence of both automatic and social reinforcement, then a lower frequency of vocalizations, particularly canonical babbling, may lead to fewer opportunities for speech to be reinforced, thus leading to further delays. Therefore, it is important to identify procedures for increasing the frequency of vocalizations in children with IDD.

Underlying Behavioral Mechanisms for Speech Development

Nonbehavioral perspectives of language development (e.g., cognitive psychology, Chomsky) suggests that language develops due to internal structures of the organism and defines language based on form and structure (Chomsky, 1959). From a behavioral perspective, language is an observable behavior that is sensitive to environmental variables (i.e., contingencies of reinforcement and punishment) that shape the particular form and function of vocal behavior. Thus, language is defined by its function, not its form (Skinner, 1957) and understanding these environmental contingencies is important for understanding language acquisition, and ultimately facilitating language development when language development is delayed (Petursdottir & Mellor, 2017). Four behavioral mechanisms are important for understanding language development. These are respondent conditioning (Rescorla, 1988), automatic reinforcement (Vaughan & Michael, 1982), social positive reinforcement (Skinner, 1969), and stimulus control (Dinsmoor, 1995a,b). These four behavioral mechanisms are described in detail below.

Respondent conditioning. When infants are born, most behavior is likely the result of reflexes. Reflexes are unconditioned responses to antecedent stimuli. That is, some antecedent external stimulation elicits the behavior (e.g., touching the lips of an infant elicits sucking; Bijou & Baer, 1965). Typically developing children, as well as children with hearing impairments, develop prespeech (e.g., crying, cooing), suggesting prespeech begins as a reflex (Schlinger, 1995). Reflexive behavior includes visible motor movements (e.g., blinking, sucking, startle movements of arms or legs), as well gross movements of the muscular system (e.g., movement of the vocal cords, tongue, oral cavity) which produces vocal sounds (e.g., cooing). A variety of sounds may be produced from gross movements of this muscular system, but perhaps the most commonly identified sound is crying. Skinner (1957) describes crying as a reflexive behavior in that it most likely first occurs as an unconditioned response to a painful situation (e.g., hunger).

After a reflex occurs, it has some effect on the environment and, in turn, the environment has an effect on the behavior (Skinner, 1957). Thus, reflexive behavior is likely to become respondent behavior. Respondent behavior is a conditioned response that is elicited by antecedent stimuli. Respondent behavior differs from operant behavior, which is maintained by consequences that have previously followed that behavior. Respondent behavior develops though respondent conditioning (Skinner, 1938).

Respondent conditioning is the behavioral mechanism by which new, previously neutral stimuli acquire the ability to elicit respondent behavior. There are five main components of respondent conditioning: the unconditioned response, the unconditioned stimulus, the neutral stimulus, the conditioned stimulus, and the conditioned response. The unconditioned response is the reflex before any conditioning occurs. The unconditioned stimulus is the stimulus or stimuli that elicit the reflex before any conditioning occurs. The neutral stimulus has no effect on behavior before conditioning occurs. The neutral stimulus becomes the conditioned stimulus through several pairings of the neutral stimulus with the unconditioned stimulus. The conditioned response is the same as the unconditioned response except that it now occurs in the presence of the conditioned stimulus instead of only in the presence of the unconditioned stimulus (Skinner, 1938). For example, when food (unconditioned stimulus) is presented to an organism, that organism will salivate (unconditioned response). If a whistle (neutral stimulus) is paired with the presentation of food (unconditioned stimulus), the whistle will become the conditioned stimulus through several pairings such that when the whistle is presented without the presentation of food, salivation occurs (conditioned response).

With respect to vocalizations, it is likely that prespeech acquires respondent behavior properties through conditioning (Schlinger, 1995). For example, crying may first occur as a reflex, elicited by hunger pains. However, a caregiver likely responds to crying with primary reinforcers such as food. Through pairing, the caregiver becomes a conditioned stimulus for the delivery of food. Thus, crying may be elicited by the presence of the caregiver, in addition to hunger pains, suggesting crying is now a respondent behavior.

Automatic reinforcement. Just as reflexes may become respondent behavior, Bijou and Baer (1965) suggested that respondent behavior may become operant. Operant behavior is behavior maintained by consequences that have previously followed that behavior, including reinforcement that occurs independent of the mediation of others. When reinforcement occurs independent of the mediation of others, it is referred to as automatic reinforcement (Skinner, 1969). For example, a person who twirls their own hair may engage in this behavior because it is automatically reinforcing. That is, the response product, or feeling of twirling his or her own hair, reinforces hair twirling independent of the mediation of others. The response product that automatically reinforces the behavior may be in the form of an unconditioned reinforcer or it may be conditioned. That is, it is possible for a neutral stimulus to be paired with other reinforcers until the neutral stimulus itself becomes a conditioned reinforcer. For example, it is possible that twirling one's own hair may start as a neutral stimulus, but be paired with a caregiver twirling one's hair (i.e., a social reinforcer) such that the response product becomes a conditioned automatic reinforcer, and now twirling one's own hair is automatically reinforcing.

Caregivers naturally pair infant speech sounds with social reinforcers (e.g., physical or verbal attention, food, warmth). While caregivers deliver these reinforcers, they likely also engage in speech sounds. In fact, research suggests that mothers tend to respond to vocal sounds of their infants with more vocal responses (e.g., cooing, gentle speech) as compared to interactive responses, such as smiles or gazes (Gros-Louis, West, Goldstein, & King, 2006). Bijou and Baer (1965) suggested that when caregivers pair their vocalizations with the delivery of reinforcers, the caregiver's speech sounds begin to acquire reinforcing properties for the infants. Thus, infants may attempt to evoke these conditioned speech sounds from others by babbling, crying, or seeking attention in other ways. Or, infants may make similar sounds themselves because these sounds are now conditioned reinforcers. It may be less effortful and more consistent for infants to emit the sounds than to evoke these sounds from others. Thus, infants' vocalizations come to resemble those of their caregivers (e.g., Schlinger, 1995; Petursdottir & Lepper, 2015). As infants hear their own vocalizations, the vocalizations may begin to develop automatically reinforcing properties because these vocalizations are conditioned reinforcers. That is, hearing their own vocalizations strengthens the future likelihood of vocalizations occurring (Bijou & Baer, 1965).

Social reinforcement. In addition to automatic reinforcement, operant behavior also includes behavior maintained by social reinforcement. When infant vocalizations begin to resemble the speech sounds of caregivers, the caregivers are likely to respond more consistently with potential reinforcers (e.g., attention), and vocalizations are likely to come under control of social positive reinforcers. Over time, caregivers are likely to deliver differential reinforcement for vocalizations that better match recognizable vocalizations (Bijou & Baer, 1965). For example, the first few times an infant says "Ahh," caregivers are likely to react with physical or vocal attention or facial expressions, all of which may reinforce vocalizations. However, over time the caregivers' reactions to the same speech sound are likely to diminish because these speech sounds are no longer novel to the caregiver. That is, caregivers may delay the delivery of attention, decrease the magnitude of attention (i.e., deliver for short periods of time), or decrease the frequency with which attention is delivered (i.e., intermittent delivery of attention). Then, when the infant says "Da" for the first time, caregivers may increase their attention again (e.g., immediate delivery, longer delivery, continuous delivery) because this is a new sound and because it more closely resembles a recognizable word (i.e., dad). Bijou and Baer (1965) referred to this as *syllabic babbling*, which is babbling that resembles the syllables that caregivers frequently emit. Additionally, at this stage of development, infants begin to echo the caregiver's sounds and chain sequences of syllables together. Often, the chained syllabic babbling begins to resemble recognizable words (e.g., "dada"). Thus, caregivers begin to differentially reinforce chains of syllabic babbling in the presence of environmental variables in which such babbling could be functional. For example, an infant may chain the sounds "ki" and "di" together to resemble the word "kitty." Caregivers may begin to reinforce such sounds in the presence of cats or pictures of cats and not in the presence of other environmental variables. In this way, infants begin to label items in their environment and learn to use these labels to request desired items (Bijou & Baer, 1965). Subsequently, functional language may be quickly learned as functional sounds come under stimulus control of relevant environmental variables (e.g., saying "kitty" in the presence of a cat) and expand to create functional units of language (e.g., words and sentences) and eventually extensive functional communication (e.g., conversations; Skinner, 1957).

Stimulus control. Social reinforcement is likely to be most effective when stimulus control is involved. Stimulus control is part of the operant conditioning process. Stimulus control refers to the behavioral mechanism in which a functional relation between a behavior and a certain antecedent condition is established. If a behavior occurs more frequently in the presence of a certain stimulus than it does in the absence of that stimulus, that behavior is said to be under stimulus control. The stimulus that evokes the behavior is called the discriminative stimulus (S^D). The S^D evokes behavior because it is a signal for the availability of reinforcement. The absence of the stimulus that evokes the behavior is a condition called stimulus delta (S^{Δ}). Thus, a behavior is more likely to occur in the presence of the S^D than an S^{Δ}. When behavior occurs in the presence of the S^D, the behavior is said to be under stimulus control (Dinsmoor, 1995a). For example, people answer the phone when the phone rings because the ringing is an S^D for reinforcement (i.e., talking to another person). People do not answer the phone when it does not ring because it is unlikely to be reinforced.

Stimulus control can be used to increase a behavior without directly delivering reinforcement contingent on that behavior, if reinforcement has been delivered in the presence of that stimulus in the past. That is, if an organism is taught that one stimulus is an S^D while another stimulus is an S^{Δ} through the process of discrimination training, the stimulus that is an S^D may become a conditioned reinforcer itself. This process may explain why certain vocalizations become conditioned reinforcers (Lepper et al., 2013). For example, a child may learn that vocalizations by caregivers are discriminative for delivery of reinforcement when that child engages in some behavior. The child may begin to emit vocalizations that match those of the caregiver because those vocalizations have been discriminative of reinforcement in the past and now function as a conditioned reinforcer for the child.

Stimulus control resembles respondent conditioning in that both evoke the occurrence of a behavior. However, there is a distinction between the two. In respondent conditioning, several pairings of an unconditioned reinforcer with a neutral stimulus will create a conditioned stimulus. This conditioned stimulus can be used to evoke the occurrence of a respondent behavior by pairing the conditioned stimulus several times with a respondent behavior. Stimulus control, on the other hand, occurs through operant conditioning. That is, a behavior is more likely to occur in the presence of some stimulus (i.e., S^D) because a specific stimulus change (i.e., the delivery of a reinforcer) has consistently followed the behavior. Thus, the S^D acquires stimulus control over the behavior. The key difference, then, between respondent conditioning and operant conditioning is the manner in which the conditioned stimulus in respondent conditioning and the S^D in operant conditioning acquire their controlling function. In respondent conditioning, the conditioned stimulus acquires its controlling function through repeated pairings of some antecedent stimulus that elicits a behavior (Rescorla, 1988). In operant conditioning, the S^D acquires its controlling function through repeated pairings with some stimulus change that signals the availability of reinforcement should a specific behavior occur (Dinsmoor, 1995a).

Respondent conditioning, automatic reinforcement, social reinforcement, and stimulus control make up four behavioral mechanisms that underlie the behavioral theory of language development (e.g., Bijou & Baer, 1965; Skinner 1957). Understanding these behavioral mechanisms may lead to effective procedures for increasing vocalizations in children with IDD.

Strategies for Increasing Vocalizations

Clinicians and researchers have evaluated several strategies for increasing vocalizations for children with IDD. These include echoic training (e.g., Stock et al., 2008), mand training (e.g., Hall & Sundberg, 1987), motor imitation training (e.g., Ross & Greer 2003), and differential reinforcement of vocalizations (e.g., Lovaas et al., 1966). However, there are several language acquisition barriers that may be present when using these strategies for increasing vocalizations. First, echoic and mand training sometimes fail to establish a vocal repertoire because the therapist is unable to prompt the child to match his or her vocal sounds, increasing the frequency of failures for the child. Increased failures may lead therapists to increase opportunities for echoing to occur; however, this increase in opportunities is often associated with an even greater increase in failures, thus creating an aversive teaching environment (Ward et al., 2007). Second, motor imitation training relies heavily on contrived contingencies (e.g., delivery of edibles or preferred toys) as opposed to naturally occurring contingencies, which is problematic when attempting to generate functional communication. Naturally occurring contingencies generate more functional behavior and are more likely to maintain behavior (Skinner, 1982). So, although motor imitation training allows the therapist to ensure the child contacts reinforcement (unlike echoic training), the child may not maintain the skills learned and may not learn functional imitation. Finally, differential reinforcement of vocalizations assumes that a child emits some vocalizations that can be shaped. However, children with IDD sometimes exhibit a complete absence of vocal language (Kjelgaard & Tager-Flusberg, 2001). In these cases, reinforcing all vocalizations may not be an effective procedure for increasing vocalizations.

Conditioning procedures. Another strategy for increasing vocalizations is to utilize conditioning procedures. Echoic training, motor imitation, and mand training all rely on social reinforcement. Children with IDD are often insensitive to social reinforcement (Petursdottir & Lepper 2015). Additionally, reinforcing the occurrence of all vocalizations assumes the child emits a variety of sounds to reinforce. Conditioning procedures do not rely on social reinforcement and do not require a child to emit vocalizations. Therefore, conditioning procedures may be more effective for increasing vocalizations for children with a minimal echoic repertoire or who exhibit minimal variation in vocalizations.

During conditioning procedures, the therapist does not deliver the reinforcer contingent on the child emitting the vocalization. Instead, reinforcers are paired with the therapist's vocalization such that vocalizations may become conditioned automatic reinforcers through respondent conditioning. After a vocalization is automatically reinforcing, it may be more likely to maintain because automatic reinforcement occurs independent of a social mediator (Skinner, 1969). Smith, Michael, and Sundberg (1996) evaluated the effects of three different pairing procedures on infant vocal behavior for two typically developing infants, ages 11 months and 14 months. Specifically, Smith, Michael, and Sundberg (1996) evaluated positive pairing (i.e., the adult stated a target sound and then delivered a preferred item – either toys or attention), neutral pairing (i.e., the adult stated a target sound, but no programmed changes to the environment followed), and negative pairing (i.e., the adult stated a target sound and then delivered an established form of punishment – specifically, the verbal reprimand, "bad girl."). Overall, the frequency of the specific target sound emitted by the infants as well as the frequency of other vocalizations increased during the positive pairing phase. There was no change in the target sound emitted by the infants during the neutral pairing, but other vocalizations increased. Finally, there was an immediate decrease in all vocal behavior during the negative pairing. These results suggest it may be possible to influence infant vocalizations through differential pairing procedures. This suggests specifically, positive pairing procedures may be an appropriate strategy for teaching vocalizations. Currently, research suggests there are three pairing procedures that may increase vocalizations: stimulus-stimulus pairing, response-stimulus pairing, and operant discrimination training.

Stimulus-stimulus pairing (SSP). SSP (also called response-independent pairing) consists of a therapist emitting a speech sound and delivering a reinforcer. That is, SSP is a response-independent relationship and as such, establishes speech sounds as automatic reinforcers by pairing the speech sound emitted by the therapist and the delivery of a reinforcer (e.g., Smith et al., 1996; Sundberg e al., 1996; Yoon & Bennett, 2000; Miguel, Carr, & Michael, 2002). Researchers have been successfully using SSP procedures to establish neutral stimuli (e.g., a teacher's voice), in general, as conditioned reinforcers for several years (e.g., Greer, Pistoljevic, Cahill, & Du, 2011).

Researchers have also used SSP procedures to establish vocalizations (e.g., nonsense syllables, whole words) as conditioned automatic reinforcers with individuals of varying ages, developmental abilities, and vocal repertoires. Overall, 12 studies have directly evaluated the effects of SSP on establishing vocalizations as automatic reinforcers (i.e., Carroll & Klatt, 2008; Esch, Carr, & Grow, 2009; Esch, Carr, & Michael, 2005; Lepper et al., 2013; Miguel et al., 2002; Miliotis et al., 2012; Norman & Knoll, 2006; Rader et al., 2014; Stock et al., 2008; Sundberg et al., 1996; Yoon & Bennett, 2000; Yoon & Feliciano, 2007). Additionally, one study evaluated the effects of SSP in conjunction with direct reinforcement on establishing vocalizations as automatic reinforcers (i.e., Ward et al., 2007). Finally, five studies (including four of the studies that directly evaluated the effects of SSP) have evaluated specific procedural variables that may influence the effectiveness of SSP (i.e., Esch, et al., 2009; Miliotis et al., 2012; Petursdottir, Carp, Matthies, & Esch, 2011; Rader et al., 2014; Sundberg, et al., 1996).

Sundberg, Michael, Partington, and Sundberg (1996) were the first researchers to evaluate SSP as a procedure for establishing vocalizations as automatic reinforcers. This study consisted of two experiments. The purpose of Study 1 was to test the use of SSP for increasing novel words or phrases. The purpose of Study 2 was to further investigate variables that may increase or decrease the effectiveness of SSP. Study 1 included five children, ages 2-4 years old. Four children exhibited severe to moderate language delays; one child did not exhibit a language delay. SSP consisted of 1-2 min of pairing (i.e., the therapist approached the child, emitted the target word or phrase, and immediately followed this with a delivery of the reinforcer – physical attention) with approximately 15 pairings per min. Data were collected on the frequency of target words or phrases emitted during 5-min pre- and post- pairing phases during which no therapist attention was delivered and no toy items were present. Four of the five children emitted more novel words or phrases during the post-pairing phase as compared to the pre-pairing phase suggesting that SSP may be appropriate for children with a variety of language abilities. Study 2 included one participant from Study 1 and attempted to answer three questions. First, the researchers evaluated why vocalizations only increased after some pairings, but not others. Specifically, they tested the effects of conducting pairings while participants exhibited various emotional states (i.e., participants who were quiet and sullen as compared to active). Results were unclear. That is, the target vocalization increased when SSP followed a period in which the participant was quiet and sullen. However, the increase in vocalizations was delayed.

Second, the researchers evaluated how long pairing effects would last in the absence of further pairings. Specifically, they extended the duration of the pre-pairing and post-pairing conditions until target vocalizations ceased to occur for a given session. Results suggested effects of SSP were robust for approximately the first 5 min following pairing, but then decreased drastically. This may suggest that SSP increases vocalizations, but only temporarily.

Finally, the researchers evaluated the possibility of altering the topography of previously paired vocalizations by pairing similar vocalizations. Specifically, in Study 1, the researchers observed that participants emitted previously paired words when the researchers introduced new words. Additionally, sometimes participants blended previously paired words with the new words when the researchers introduced new words. Therefore, for Study 2, the researchers purposefully introduced new words that resembled a previously paired word. No blending of the new word with the previously paired word was observed, suggesting this procedure did not result in better pairing effects.

Overall, Sundberg et al.'s (1996) results suggested SSP may increase novel vocalizations by pairing neutral words or phrases with reinforcers, specifically physical attention. It is important to note that the participants acquired these novel vocalizations without the use of direct reinforcement for participant vocalizations or direct echoic training, suggesting vocalizations may have functioned as conditioned automatic reinforcers for these participants. This is important because it suggests SSP may be an effective procedure for increasing vocalizations.

Although some studies have suggested the efficacy of SSP, across all studies evaluating the effectiveness of SSP for increasing vocalizations, researchers report mixed results. Overall, seven of the 12 studies that directly evaluated the effects of SSP suggested that SSP established speech sounds as automatic reinforcers for *all* participants (i.e., Esch et al., 2009; Lepper et al., 2013; Miliotis et al., 2012; Rader et al., 2014; Sundberg et al.,1996; Yoon & Bennett, 2000; Yoon & Feliciano, 2007). Two of these same twelve studies suggested SSP established speech sounds as automatic reinforcers for *some* participants (i.e., Carroll & Klatt, 2008; Miguel et al., 2002). Three of these twelve studies suggested SSP established speech sounds as automatic reinforcers for *none* of the participants (i.e., Esch et al., 2005; Normand & Knoll, 2006; Stock et al., 2008). Finally, the one study that evaluated the effects of SSP in conjunction with direct reinforcement suggested SSP and direct reinforcement established speech sounds as automatic reinforcers for *all* participants (i.e., Ward et al., 2007).

Across all studies evaluating SSP, there are several variations that might alter the effectiveness of this pairing procedure. In 2015, Shillingsburg, Hollander, Yosick, Bowen, and Muskat published a review of all studies that used SSP to increase vocalizations in children with language. The authors analyzed results of 13 studies with respect to participant characteristics and procedural variations. Results were categorized as weak, moderate, or strong treatment effects. Treatment effects were quantified by calculating the nonoverlap of data points (i.e., fewer treatment data points overlapping with control data points resulted in a stronger treatment effect). Overall, Shillingsburg et al. (2015) reported a moderate treatment effect across studies, with 34% of evaluations showing a weak effect, 49% of evaluations showing a moderate effect, and 17% of evaluations showing a strong effect. With respect to participant characteristic differences and procedural differences, it seems that SSP may be best suited for young individuals with language delays. Additionally, SSP procedures should include delay conditioning, edibles as preferred items, and a control for adventitious reinforcement. However, Shillingsburg et al. clearly state that limitations exist in this review and that further research is needed to clearly identify the best population with which to use SSP and the most effective SSP procedures.

Five studies have specifically evaluated procedural variables that may be correlated with more effective SSP procedures. For example, Esch, Carr, and Grow (2009) assessed several variations in procedural variables together on the effects of SSP. Specifically, Esch et al. added a nontarget sound trial, an observing prompt, an exaggerated prosodic pattern (known as *motherese*; Falk, 2004), and varied inter-trial intervals. The authors coined these procedures as "enhanced" SSP procedures. Esch et al. hypothesized the addition of a nontarget sound trial would enhance the effects of pairing because it may increase the saliency of the reinforcement contingency as a preferred item consistently follows one sound while a preferred item never follows another sound. Additionally, they hypothesized that the addition of an observing prompt (i.e., stating, "Look" and waiting until the participant oriented his or her eyes toward the therapist) would increase the likelihood that the participant was attending to the therapist when the SSP presentation occurred. The researchers also suggested the addition of motherese would increase the likelihood that the participant would discriminate between speech sounds emitted by the therapist during the session and speech sounds emitted by the therapist during the ITI. Finally, they hypothesized that the use of varied ITIs would ensure SSP effects were attributable to the conditioning procedure and not because the participant could predict when reinforcement was available. Overall, the researchers reported an increase in vocalizations as compared to the nontarget vocalizations for the three participants for whom enhanced SSP procedures were used. Unfortunately, the results are limited because the enhanced procedures were not evaluated individually making the effects of any one procedural variable unclear. Additionally, the researchers did not attempt SSP procedures without the enhanced procedures first with these participants. Because mixed results of SSP have been observed across studies, it is plausible that these enhanced SSP procedures may have been necessary to increase vocalizations for these participants.

Overall, research seems to suggest SSP may be appropriate for increasing vocalizations for children with limited vocal repertoires. However, findings are limited and inconsistent. Additionally, there are numerous procedural variations between studies that make it difficult to identify the most effective combination of procedures. As such, researchers have begun to evaluate other procedures for increasing vocalizations that continue to utilize conditioning procedures.

Operant discrimination training (ODT). A second procedure that has been suggested to establish neutral stimuli as conditioned reinforcers is ODT. ODT consists of a therapist delivering a reinforcer contingent on an arbitrary operant response (e.g., button press) in the presence of a discriminative stimulus (i.e., when a target sound is emitted by the therapist) and not delivering a reinforcer in the absence of this discriminative stimulus (i.e., when a nontarget sound is emitted by the therapist). With respect to speech sounds, ODT conditions one speech sound emitted by the therapist as a discriminative stimulus for reinforcement by delivering a reinforcer contingent on an arbitrary operant response following the presentation of a sound by the therapist (e.g., Lepper et al., 2013).

Like SSP, ODT has been used for several years (e.g., Lovaas et al., 1996). However, of particular note was a study conducted by Taylor-Santa, Sidener, Carr, and Reeve (2014). They evaluated the optimal methods for establishing conditioned reinforcers using ODT. Unlike previous research, these authors included an S^{Δ} condition because they hypothesized that the inclusion of an S^{Δ} trial may result in faster discrimination because it may also increase the saliency of the reinforcement contingency. Pre- and post-test conditions occurred in which the rate of switch presses that were followed by the S^D was compared to the rate of switch presses that were followed by the S^{Δ}. Discrimination training consisted of two steps. In the first step, the research presented the S^D, then presented the switch press. If the participant did not press the switch within 3 s, the therapist prompted the participant to engage in the response. Following a switch press (either independent or prompted), the therapist delivered a reinforcer. Discrimination training continued until the participant engaged in 100% correct responding across two consecutive sessions across two days. In the second step, the researcher presented the S^{Δ}. Sessions consisted of 20 trials (i.e., 10 S^D trials and 10 S^{Δ} trials presented in a quasi-random order). The researcher either presented the S^D or the S^{Δ}, then the researcher presented the switch press. If the researcher presented the S^D and the participant did not press the switch within 3 s, the therapist prompted the participant to engage in the response. Following a switch press (either independent or prompted), the researcher delivered a reinforcer. If the researcher presented the S^{Δ} and the participant pressed the switch, the trial was terminated. If the researcher presented the

S^{Δ} and the participant did not press the switch for 4 s, the trial was terminated. This continued until 100% correct responding was observed across two days. Then, post-tests were conducted in which the rate of switch presses that were followed by the S^{D} was compared to the rate of switch presses that were followed by the S^{Δ}. Results suggested discrimination training was successful for all three participants. That is, all three participants pressed the switch more often in the presence of the S^{D} than in the presence of the S^{Δ}. All participants met the 100% correct responding during the second step of the discrimination training phase within six to nine trials. Additionally, results suggested discrimination training successfully conditioned the S^{D} as a reinforcer. That is, the participants pressed the switch at a higher rate when switch pressing was followed by the S^{D} than when switch pressing was followed by the S^{Δ}. These results suggest that discrimination training can occur quickly and may successfully condition a neutral stimulus as a reinforcer. This study further suggests similar procedures may be used to condition vocalizations as reinforcers, and ultimately increase vocalizations for children with minimal vocal repertoires.

To date, only one study has directly evaluated the effects of ODT on establishing vocalizations as automatic reinforcers. Lepper, Petursdottir, and Esch (2013) compared the effects of ODT to SSP on increasing vocalizations with three children ages 2-4 years, all diagnosed with ASD. The researchers used an alternating-treatments design to compare ODT, SSP, and a control condition. During all conditions, researchers presented target and nontarget trials for 20 trials each. The researcher presented the sound three times per trial and conducted an ITI of 10-15 s. Additionally, the researcher began each trial with an observation response (i.e., the therapist stated, “Look”). The researchers conducted all three conditions with each participant twice using two different sets of sounds. During the ODT condition, the target sound served as the S^{D} and the nontarget sound served as the S^{Δ}. Arm raising served as the task that would result in reinforcement in the presence of the S^D only. When the therapist presented the target sound and an arm raise followed, then the therapist delivered the preferred item. If an arm raise did not follow, then the therapist prompted the participant to raise his or her arm, and then delivered the preferred item. When the therapist presented the nontarget sound, and an arm raise followed, the therapist ignored it. If an arm raise did not follow, then the therapist began the ITI. During the SSP condition, the researchers used SSP procedures similar to those in Esch et al. (2009). That is, after the therapist stated the target sound, he or she delivered the preferred item simultaneously with the third and last presentation of the sound. The ITI began following consumption of the preferred item. After the therapist stated the nontarget sound, the therapist began the ITI. During the control condition, procedures were identical to the SSP condition except that the therapist inserted a 20 s delay before the delivery of the preferred item following the presentation of the target sound. Separate from the comparison of SSP and ODT on increasing vocalizations, Lepper et al. also conducted a preference evaluation in which each condition was paired with a different colored paper. At the start of each preference session, the participants chose a colored paper and immediately experienced four trials from the relevant condition. Overall, SSP and ODT were equally effective and resulted in an increase in target sounds for all participants (except for one target sound for one participant). In addition, participants more often preferred the ODT condition. These results provide preliminary evidence that ODT may be an appropriate strategy for increasing vocalizations.

Overall, research suggests that ODT may be appropriate for increasing vocalizations for children with limited vocal repertoires. However, findings are limited, with only one study evaluating the effects of ODT on increasing vocalizations (Lepper et al., 2013). Additionally, this study did not necessarily produce better results than SSP. As such, further research into the effectives of ODT is necessary as well as further research into other procedures for increasing vocalizations that may be more consistent than SSP and ODT.

Response-stimulus pairing (RSP). RSP (also called response-contingent pairing) consists of a participant initiating a stimulus-stimulus pairing by first completing an arbitrary operant response (e.g., button press). That is, RSP is a response-contingent relationship because the pairing is contingent upon an arbitrary operant response completed by the participant. Like the other two conditioning procedures, RSP may effectively increase vocalizations by conditioning the therapist's vocalizations as automatic reinforcers (e.g., Lepper, 2014).

Several researchers have used RSP to condition neutral stimuli as reinforcers (e.g., Dorow, 1980; Dozier, Iwata, Thomason-Sassi, Worsdell, & Wilson, 2012; Sidowski, Kass, & Wilson, 1965). To date, only one study has been published that directly evaluates the effects of RSP on increasing vocalizations (i.e., Lepper and Petursdottir, 2017). This was a two-part study. The purpose of Study 1 was to compare the effects of RSP to SSP on increasing vocalizations. The purpose of Study 2 was to increase the rate of the vocalizations established in Study 1 using differential reinforcement. Three boys with ASD and severely delayed speech participated in both studies. During Study 1, the researchers identified four vocalizations per participant by observing speech sounds already exhibited by each participant and creating novel vocalizations made up of these already existing vocalizations (e.g., if a participant was observed to say "moo" and "bee," a novel vocalization may be "boo"). Both RSP and SSP sessions consisted of 20 trials (i.e., 10 target sound trials and 10 nontarget sound trials). During SSP target trials, the researcher presented the target sound three times and delivered a preferred item simultaneously with the third presentation of the target sound. During RSP target trials, the researcher presented the target sound and preferred item in the same way, but only after the participant pressed a button. During SSP nontarget trials, the therapist simply presented the nontarget sound three times. During RSP nontarget trials, the therapist only presented the nontarget sound after the participant pressed a button. During SSP, the researchers presented the target and nontarget sounds at a rate yoked to that of the RSP condition.

Overall, all participants demonstrated a higher rate of target vocalizations during the RSP condition as compared to the rate of target vocalizations during the SSP condition. Lepper suggested three reasons why this might be the case. First, it is possible that the RSP condition allows for the preferred item to be delivered exactly when it is desired by the participant. That is, the participant must engage in the response first and, as such, can control when the reinforcer is delivered. In contrast, the reinforcer is delivered on a set time schedule during SSP. Second, it is possible that the contrafreeloading effect (Inglis, Forkman, & Lazarus, 1997) played a role in RSP. That is, Inglis et al. (1997) suggested that greater effects are often demonstrated when participants are required to work for a preferred item. This also aligns with research suggesting that children prefer contingent to noncontingent delivery of preferred stimuli (Luczynski & Hanley, 2009). Third, the button press may have functioned as an observing response, which previous research has suggested may ensure the participant attends to the auditory stimulus (e.g., Esch et al., 2009). Although Lepper included an observing response (i.e., “Look”) in the SSP condition, it is possible that the button press made it more likely that the participant attended to the appropriate stimuli (i.e., the auditory stimulus and preferred stimulus pairing) as opposed to other stimuli (e.g., therapist mouth movement).

During Study 2, Lepper (2017) attempted to increase the rate of vocalizations established in the RSP condition of Study 1 using direct reinforcement. Specifically, RSP was conducted identically to Study 1 except that all target vocalizations resulted in the delivery of the preferred item. Additionally, if the participant emitted the target vocalization during the ITI, the ITI was reset and the preferred item was delivered. Finally, if the target vocalization was emitted during the consumption period, then the period was extended for an additional 15 s access to the preferred tangible item or the participant was offered an additional bite of the preferred edible, allowing the participant to access the preferred item for a longer period of time. Overall, all three participants demonstrated an increase in target sounds during the direct reinforcement phase. This increase was higher than baseline levels in Study 2 and RSP levels in Study 1. Additionally, an increase in target sounds was replicated across two target sounds for each participant.

Overall, research seems to suggest that RSP may also be appropriate for establishing vocalizations as automatic reinforcers as a way of increasing vocalizations for children with limited vocal repertoires. Additionally, there is preliminary evidence to suggest that RSP may be more effective than other procedures (at least, SSP). However, findings are limited. As such, further research into the effectiveness of SSP, ODT, and RSP is necessary. Additionally, research should be conducted to compare SSP, ODT, and RSP with the same participant to determine whether one is more effective than the others for increasing vocalizations. Conducting all three conditioning procedures with the same participant helps control for characteristic differences across participants. However, further research should also specifically evaluate participant characteristic variations that may alter the effectiveness of these conditioning procedures (e.g., articulation skills, attending skills). This may allow researchers to identify effective and efficient procedures for increasing vocalizations, and ultimately increasing communication, when speech development is delayed.

Purpose

The purpose of the current study was to evaluate the effects of three conditioning procedures suggested in the current literature to condition speech sounds as automatic reinforcers. The purpose of Study 1 was to compare the effects of SSP, RSP, and ODT with each participant on the frequency of novel vocalizations exhibited by individuals with a minimal vocal repertoire. Participants for whom all three conditioning procedures were unsuccessful at increasing the frequency of novel vocalizations participated in Study 2. The purpose of Study 2 was to evaluate three potential variables that may alter the effectiveness of SSP, RSP, and ODT. Specifically, articulation, reinforcer efficacy, and attending were evaluated as three potential variables that may change the effectiveness of these procedures.

General Method

Participants

Seven participants, ages 1 to 3, with a minimal or emerging echoic repertoire participated (see Table 1 & 3 for participant demographics). A minimal echoic repertoire was defined as a score of five or less on the Early Echoic Skills Assessment (EESA; Esch, 2008) and an emerging echoic repertoire was defined as a score between 15 and 25. The EESA was conducted with each participant prior to the start of the study (see Table 2 for pre-conditioning EESA scores). A score of five or less was selected as a minimal echoic repertoire because individuals with this score may echo, at most, five unique vowel or consonant sounds. However, they are unable to echo sounds with more than one syllable. This score suggests that these participants lack an echoic repertoire. A score between 15 and 25 was selected as an emerging echoic repertoire because individuals with this score echo several vowel or consonant sounds, but are still unable to echo sounds with more than one syllable. Participants included individuals with IDD (e.g., ASD, Down syndrome), as well as typically developing individuals (i.e., individuals with no known diagnosis), recruited from the Edna A. Hill Child Development Center (CDC) at the University of Kansas.

Setting and Materials

Trained therapists conducted sessions in private session rooms (3 m by 3 m) adjacent to the individual's classroom. The session rooms were equipped with one-way mirrors and were barren except for condition-specific stimuli (i.e., relevant task materials, relevant preferred edibles). The walls of each room had been painted different colors (i.e., purple, green, yellow), and the rooms were correlated with a specific experimental condition facilitate discrimination between experimental conditions. Additionally, therapists conducting sessions wore shirts that corresponded in color with the rooms. Therapists also wore gloves and carried a pen, timer, and a clipboard with a data sheet to facilitate trial accuracy (i.e., order of trials, total number of trials, and duration of trials). Only a therapist and the participant were present in the session room. Data collectors observed from behind the one-way mirror.

Periodically, generalization probes were conducted in both novel session rooms and in the participant's classroom. During generalization probes in the session room, items that were typically present in the individual's classroom were available (e.g., leisure items, manipulatives, toys, books). Only a therapist and the participant were present in the session room. Data collectors observed from behind the one-way mirror. During generalization probes in the individual's classroom, no programmed changes occurred to the environment. Generalization probes were conducted in the classroom during free-play periods. Generally, leisure items were available. Classroom teachers and other students were present and might have interacted with the participant. A data collector was also present to videotape and stayed in close proximity to the participant, while being as discreet as possible.

Stimuli. Two of six task responses were assigned to each participant based on the outcome of the task assessment (described below). For six of the seven participants, possible task responses included a *button press* task, a *card touch* task, a *Block-in-Bucket* task, a *Stack Card* task, a *Mail Letter* task, and a *String Beads* task. The *button press* task consisted of a BIGmack® that created an audible click when pressed, but no programmed noise. The *card touch* task consisted of a laminated square card that was relatively the same size as the BIGmack® button. The *Block-in-Bucket* task consisted of a single Lego Duplo and a plastic cup that was just larger than the Lego Duplo. The *Stack Card* task consisted of a single laminated circular card with a hole in the middle and a dowel rod that was just skinnier than the hole in the circular card. The *Mail Letter* task consisted of a single laminated rectangular card and a tissue box with a slot in the top just wider than the laminated card. Finally, the *String Beads* task consisted of a shoe string and a single large bead.

For one participant (Maisy), the possible task responses were modified because she lacked the pre-requisite skills to complete the tasks described above. For Maisy, possible task responses included a *Drum* task, a *Tambourine* task, a *Piano* task, a *Bells* task, a *Stick* task, and a *Ring* task. The *Drum* task consisted of a small toy hand drum. The *Tambourine* task consisted of a small toy tambourine about the same size as the drum. The *Piano* task consisted of a small, 10-key piano. The *Bells* task consisted of toy jingle bells with a handle. The S*tick* task consisted of a drum stick. Finally, the *Ring* task consisted of the largest ring from a ring-stacker toy.

Study 1

Purpose and Experimental Arrangement

The purpose of Study 1 was to compare the effects of SSP, RSP, and ODT on the frequency of novel vocalizations exhibited by individuals with a minimal vocal repertoire.

A multielement design combined with a multiple baseline design within and across participants and a reversal design was used to evaluate treatment effects. Sessions were conducted in blocks of three sessions (i.e., each experimental condition occurred once in a series). A block of sessions was conducted one to two times a day, three to five days a week. Experimental conditions were presented in a quasi-random order (i.e., the order of experimental conditions within a block were randomly selected by assigning each condition a number and entering these numbers into a random number generator, but the same experimental condition did not repeat more than twice in a row across blocks). All sessions consisted of 20 trials with an ITI of 10 s. For three participants (i.e., Benny, Mason, and Antonio) each trial consisted of one presentation of either the target or the nontarget sound. These were the first three participants to participate in this study and we chose one presentation of the sound based on Miliotis and colleagues (2012) findings that one sound per pairing resulted in a higher rate of vocalizations than three sounds per pairing. However, Miliotis and colleagues (2012) only included two children, and the majority of the current literature across SSP, RSP, and ODT, presents three sounds per pairing. Therefore, to better replicate the most common methodology found in the current literature, each trial consisted of three presentations of either the target or the nontarget sound for the remaining four participants. The therapist presented each sound at a rate of one sound per second. The therapist presented the target sound during 10 of the trials and the nontarget sound during the other 10 trials. The therapist quasi-randomly alternated between target and nontarget trials (i.e., the therapist never conducted a target or nontarget trial more than twice in a row). To ensure trial orders were quasi-random, six different trial orders were created by randomly selecting either a target or nontarget trial ten times. Whichever trial type was not chosen, then occurred immediately following the chosen trial type. For example, if a target trial was randomly chosen, then a nontarget trial immediately followed and vice versa. Therapists randomly chose a number between one and six before each session, and conducted the trial order associated with the chosen number. All six trial orders had to be used before a trial order was repeated. SSP condition was always conducted in the yellow room; the RSP condition was always conducted in the purple room; and the ODT condition was always in the green room. Therapists varied across conditions and across days.

Response Measurement and Reliability

The primary dependent variable was the frequency of target sounds exhibited by the participant (reported as responses per minute). A target sound was defined as the participant emitting the full target sound correctly (i.e., the vowel and consonant were emitted in the correct order with the correct articulation and approximations were excluded). Every instance of the target sound, regardless of the length of pause between utterances, was recorded.

Secondary dependent variables included the frequency of nontarget sounds (reported as responses per minute) and the level of other sounds (reported as percentage of intervals). A nontarget sound was defined as the participant emitting the full nontarget sound correctly (i.e., the vowel and consonant were emitted in the correct order with the correct articulation and approximations were excluded). Every instance of the nontarget sound, regardless of the length of pause between utterances, was recorded. Other sounds were defined as the participant emitting any other sound (made up of consonants, vowels, or a combination of consonants and vowels) that did not directly match the target or nontarget sound. The occurrence of other sounds was recorded when any instance of other sounds occurs at any point during a 10 s interval. Partial interval data collection for other sounds was used because participant vocalizations often occurred quickly, varied in length, and consisted of several varied syllables without a pause in between. Therefore, it was difficult to define the start and stop of one vocalization, and so a partial interval measure seemed more appropriate. Negative vocalizations (e.g., crying, screaming, whining) as well as naturally occurring sounds (e.g., grunting, sighing, heavy breathing, burping, laughing) were excluded from other sounds.

Additionally, data were collected on the frequency of the task response during the RSP and ODT conditions. Specifically, a *button press* response was defined as the participant pressing the button in a downward motion with his or her hand with enough force to make an audible click (i.e., the mechanical click from depressing the button. The button was not programmed to make any additional sound). A *card touch* response was defined as the participant pressing the card in a downward motion with at least one of his or her hands. Hitting the card with two hands simultaneously was counted as one instance. A *block-in-bucket* response was defined as the participant placing the block completely into the bucket and releasing. That is, placing the block into the bucket, but pulling the block back out without releasing did not count. A *stack card* response was defined as the participant placing the hole in the card onto the dowel rod completely and releasing such that the card slides to the bottom of the dowel rod. Therefore, placing the hole in the card onto the dowel rod, but only sliding the card part way down the dowel rod before removing the card from the dowel rod did not count. A *mail letters* response was defined as the participant placing the letter into a box such that the majority of the card was in the box (i.e., a sideways card that stuck out of the top of the box still counted as a response) and releasing the letter. Therefore, placing the card in the box, but pulling the card back out without releasing did not count. Finally, a *string bead* response was defined as the participant placing the string completely through the bead and releasing the bead. Therefore, placing the string through the bead, but then pulling the bead off the string or pulling the string away from the bead did not count.

For Maisy only, a *Drum* response was defined as Maisy pressing the drum in a downward motion with at least one hand with enough force to make an audible noise. Hitting the drum with two hands simultaneously counted as a single response. A *Tambourine* response was defined as Maisy pressing the tambourine in a downward motion at least one hand with enough force to make an audible noise. Hitting the tambourine with two hands simultaneously counted as a single response. A *Piano* response was defined as Maisy pressing any key with her hand with enough force to make an audible noise. Hitting the keys with two hands simultaneously or hitting multiple keys simultaneously counted as a single response. A *Bells* response was defined as Maisy picking up the bells far enough that the bells did not touch the table. A *Stick* response was defined as Maisy picking up the stick far enough that the stick did not touch the table. Finally, a *Ring* response was defined as Maisy picking up the ring far enough that the ring did not touch the table.

The rate of target and nontarget sounds that occurred during the ITI were reported because sounds under the control of automatic reinforcement are likely to occur in the absence of socially mediating stimuli (e.g., attention from the therapist, potential delivery of preferred food). The percentage of intervals during which other sounds occurred were reported for both trial and ITI combined because some intervals partially occurred during the trial and partially during the ITI. Trained observers collected data using a computer program called BDataPro. Data collection occurred in-vivo or from video if in-vivo data collection was not possible (e.g., no observer was available, the computer malfunctioned). A second independent observer collected data independently during at least 33% of all sessions to assess interobserver agreement (IOA). When possible both observers collected data simultaneously. However, sometimes was necessary for the second observer to collect data at a different time than the first observer (e.g., two observers are not available during a session). In this case, both observers still collected data independently, but the first observer may have scored the session in-vivo, while the second observer may have scored the session later from video. Proportional reliability was calculated for all primary and secondary dependent variables except for other sounds for which partial interval reliability was calculated instead. Proportional reliability was calculated by dividing the sum of the total agreement intervals and the proportion of agreement per interval by the total number of intervals and multiplying by 100%. Partial interval reliability was calculated by dividing the number of agreement intervals by the total number of intervals and multiplying by 100%. Across baseline, conditioning, and contingency reversals, the average IOA across participants for the target sound was 98% (range: 77% to 100%), for the nontarget sound was 99% (range: 81% to 100%), for other sounds was 84% (range: 67% to 98%), and for task response was 95% (range: 58% to 100%). Across classroom and session room generalization probes, the average IOA across participants for the SSP target sound was 99% (range: 94% to 100%), for SSP nontarget sound was 100% (range: 99% to 100%), for RSP target sound was 99% (range: 94% to 100%), for RSP nontarget sound was 99% (range: 83% to 100%), for ODT target sound was 99% (range: 93% to 100%), for ODT nontarget sound was 100% (range 100% to 100%), and for other sounds was 85% (range: 53% to 96%). The average IOA scores remain above 90% for target and nontarget sounds, and above 80% for other sounds, but the ranges vary greatly, with as low as 58% for one participant. However, proportional reliability is a relatively stringent measure of reliability. Additionally, participants' vocalizations varied greatly with respect to speed and variety. Thus, with some participants it was easy to discriminate one vocalization from another. However, with other participants this was extremely difficult. This may have contributed to varying IOA.

Procedural Fidelity

Data were also collected on several other variables to assess procedural fidelity. First, data were collected on the presentation of the target sound and the nontarget sound by the therapist. The presentation of the target sound was defined as the therapist emitting the full target sound correctly (i.e., the vowel and consonant were emitted in the correct order with the correct articulation). The presentation of the nontarget sound was defined as the therapist emitting the full nontarget sound correctly (i.e., the vowel and consonant were emitted in the correct order with the correct articulation). Second, data were collected on the delivery of the preferred edible. The delivery of the preferred edible was defined as the therapist placing the programmed preferred edible into the participant's hand or mouth. Finally, data were collected on therapist prompts for the participant to engage in the task. Prompts consisted of model and physical prompts. A model prompt was defined as the therapist appropriately engaging in the task. A physical prompt was defined as the therapist physically guiding the participant to engage in the task.

The level of procedural fidelity with respect to target and nontarget sound presentation was calculated by identifying the percentage of sessions with an error of omission or an error of commission with respect to the number of target and nontarget trials. An error of omission occurred if less than 10 target or nontarget trials occurred during a session. An error of commission occurred if more than 10 target or nontarget trials occurred during a session. The level of procedural fidelity with respect to the number of correctly implemented trials was calculated by dividing the number of correctly implemented trials by the total number of trials and multiplying by 100%. A correctly implemented trial was defined as (a) the steps for a given condition were conducted in the correct order (described below), (b) the target sound is followed by a preferred edible, (c) the nontarget sound is not followed by a preferred edible, and (d) if prompting was necessary, a model prompt occurred first and was then followed by a physical prompt. Finally, the level of procedural fidelity with respect to omission contingencies (described below) was calculated by identifying the percentage of sessions during which a preferred edible was delivered within 10 s of the participant emitting the target or nontarget sound. On average, across participants, the percentage of sessions that included the correct number of target and nontarget sessions was 89% (range: 80% to 99%), the percentage of trials that were implemented correctly was 95% (range: 79% to 100%), and the percentage of trials during which the omission contingency was implemented correctly was 90% (range: 45% to 100%). The average for all aspects of procedural fidelity is at or above 89%. However, the ranges vary greatly, with as low as 45% accuracy. However, the lower scores are associated with only one participant – Benny. Benny was the first participant to participate in this study and therapists were re-trained with respect to procedural fidelity throughout and following conducting procedures with Benny. Procedural fidelity was at or above 90% for all other participants.

Experimental Conditions

Pre-assessments. Several pre-assessments were conducted prior to the start of the study. First, the EESA was conducted to evaluate the participant's echoic repertoire and determine if the participant met inclusion criterion. Second, a paired stimulus preference assessment (PSPA; Fisher et al., 1992) was conducted to identify three preferred item from which the participant could choose to be paired with speech sounds during the experimental conditions. Third, a task assessment was conducted to identify the task to be paired with speech sounds during the relevant experimental conditions. Finally, observations were conducted to identify six different speech sounds that consisted of sounds already emitted by the participant to be used during experimental conditions as the target and nontarget sounds.

Early echoic skills assessment (EESA). The EESA was conducted according to the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) guide (Appendix A; Esch, 2008). The EESA was conducted in a session room or a quiet part of the child's daily classroom to minimize distractions. A trained therapist conducted the EESA and collected data using a pencil and paper. A trained observer also collected data independently and simultaneously with the therapist using a pencil and paper. The therapist presented each sound from Level 1 of the EESA three times. Data were collected on the accuracy with which the participant imitated the vocal sound emitted by the therapist. Specifically, the participant's best response across the three presentations was recorded. No points were given if no response occurred, incorrect vowels occurred, or the response was missing syllables. A half a point was given if there was a recognizable response, but it was incorrect or missing consonants or included extra syllables. Finally, a full point was given if the correct sound was made with the correct number of syllables. Only participants who scored a five or lower or between 15 and 25 on the EESA participated in this study.

IOA was calculated for at least 33% of all trials per participant. An agreement was defined as both observers scoring the same point value for a given speech sound (i.e., zero, half, or one point). IOA was calculated by dividing the number of agreements by the sum of the number of agreements and disagreements and multiplying by 100%. On average, across participants, IOA for the EESA was 97% (range: 91% to 100%).

Paired-stimulus preference assessment (PSPA). A PSPA was conducted according to the methodology first suggested by Fisher et al. (1992). The PSPA was conducted in a session room or a quiet area of the participant's classroom to minimize distractions. Specifically, 10 food items were included that caregivers or teachers suggested the participant preferred. Items were presented to the participant in pairs and each food item was paired with every other food item once. Participants were allowed to sample each food item before pairings occurred. A trained therapist conducted the PSPA and collected data using a pencil and paper. A trained observer also collected data independently and simultaneously from the therapist using a pencil and paper. Specifically, the therapist presented two food items and said, "Pick your favorite." The participant was allowed 15 s to select a food item. If the participant selected a food item, then participant was allowed to consume that item and the other item was removed. If the participant did not select a food item within 15 s, both items were removed and represented a second time. If the participant still did not select a food item within 15 s, then both items were removed and "no selection" was recorded. Data were collected on selection (or no selection), consumption (i.e., acceptance only, acceptance and consumption, or acceptance and expulsion), and location of choice (to ensure a side bias did not exist). Data were analyzed based on frequency of selection and consumption. The three items that were selected and consumed most frequently were paired with the speech sounds during the experimental conditions (i.e., the participant chose between these three items before each experimental session and the therapist used the chosen food item for the subsequent session).

IOA was calculated for at least 33% of all trials per participant. An agreement for selection was defined as both observers recording the same selection of food item for a given trial or both observers recording no selection. An agreement for consumption was defined as both observers recording acceptance (i.e., the participant picked up the selection), consumption (i.e., the participant placed the chosen food item in his or her mouth and swallowed), and or expulsion (i.e., the participant spit out all or a portion of the chosen food item after it was in his or her mouth). Finally, an agreement for location of choice was defined as both observers recording the same location (i.e., left or right) of the chosen food item. IOA was calculated by dividing the number of agreements by the sum of the number of agreements and disagreements and multiplying by 100%. On average, across participants, IOA for selection was 99% (range: 96% to 100%), for consumption was 100% (range: 98% to 100%), and for location of choice was 99% (range: 96% to 100%).

Single-stimulus preference assessment (SSPA – Maisy only). A different assessment of preference was conducted for Maisy. Maisy's caregivers and case managers reported that she lacked the pre-requisite skills to scan an array of items based on observations and direct assessments (e.g., VB-MAPP). Additionally, Maisy was diagnosed with a cortical visual impairing, and her vision specialist indicated that Maisy may have trouble processing visual stimuli. However, through Maisy's classroom programming, Maisy had demonstrated the ability to consistently press a BIGmack® button to request more of food. Therefore, Maisy's preference for food was assessed by using an SSPA during which Maisy pressed a BIGmack® button to access a given food. The BIGmack® button was used instead of letting Maisy have free access to a given food (as is typical in an SSPA) because Maisy also lacked the ability to feed herself with utensils or with her fingers.

The SSPA for Maisy was conducted similarly to the methodology first suggested by Pace, Ivancic, Edwards, Iwata, and Page (1985). The SSPA was conducted in a session room to minimize distraction. Just as with the other participants, 10 food items that caregivers or teachers suggested that Maisy preferred were included. Items were presented to Maisy one at a time for 2 min each. Each food item was presented to Maisy three times in a quasi-random order (i.e., all 10 items were presented once before any item was presented again). Maisy was allowed to sample each food item before the assessment began. To do so, therapists stated, "Maisy, this is the (food item)" and placed a bite of food in her mouth. After Maisy swallowed this first bite, the therapist placed Maisy's hand on the BIGmack® button and stated, "Here is your button. If you want more, you can press it." A trained therapist was in the session room with Maisy and conducted the SSPA.

Two trained observers collected data independently and simultaneously using ABC Data Pro on an iPod. Data were collected on acceptance, consumption, and expulsion, all defined identically as they were defined for other participants. Data were also collected on rejection (i.e., no food passed the plane of Maisy's lips) and button press (i.e., pressing the button in a downward motion with her hand with enough force to make an audible click). Data were analyzed based on the frequency of button presses and the percentage of bites that were accepted and consumed. The three items that were selected and consumed most frequently and had a rate of responding at or above efficiency (i.e., Maisy was observed to swallow a bite about every 10 s, and so efficient responding should be a button press every 10 s - or six responses per minute) were paired with the speech sounds during the experimental conditions. Unlike other participants, one of the three preferred edibles was quasi-randomly chosen before each session (i.e., each preferred edible had to be chosen once before it could be chosen again).

IOA was calculated for at least 33% of all sessions. An agreement was defined for all target behaviors as the same number of responses for a given 10 s interval. That is, exact reliability was calculated. IOA was calculated by dividing the number of agreement intervals by the sum of the number of agreement and disagreement intervals and multiplying by 100%. On average, IOA for acceptance was 90% (range: 67% to 100%), for expulsion was 80% (range: 67% to 100%), for consumption was 81% (range: 58% to 100%), for rejection was 100% (range: 100% to 100%), and for button press was 84% (range: 67% to 100%).

Single-stimulus task assessment (SSTA). Because the tasks included in this study may vary in difficulty, an SSTA similar to the methodology first suggested by Pace, Ivancic, Edwards, Iwata, and Page (1985) for the SSPA was conducted to determine a task that the participant could complete independently. The SSTA was conducted in a session room to minimize distraction. Specifically, six task items were included in the SSTA. The same six task items were included for all participants, with the exception of Maisy who lacked the pre-requisite skills to complete the tasks used with other participants. Each task item was presented to the participant individually. Sessions were 5 min in length. Each item was presented three times in a quasi-random order (i.e., each item was assigned a number and numbers were chosen using a random number generator; all six numbers had to be chosen once before a number could repeat). Prior to the start of the session, a therapist prompted participants to complete the task once using first a verbal prompt, followed by a model prompt, and then followed by a physical prompt to ensure the participant experienced the task. At the start of the session, the therapist stated, "You can do whatever you want." During session, a therapist stayed present in the room but minimized interactions with the participant (i.e., if the participant initiated an interaction with the therapist, the therapist would respond, but the therapist did not initiate interactions). The participant was free to engage with the task. No programmed consequences occurred if the participant engaged in the task.

Two trained observers collected data independently and simultaneously using the ABC Data Pro application on an iPod. The observers remained behind the one-way mirror to collect data. Data were collected on the frequency of task completion, but reported as a partial interval measure (i.e., percentage of 10 s intervals during which at least one task completion occurred). A partial interval measure was used to evaluate task completion such that it was more likely we identified a task that the participant completed throughout a session and not several times at the beginning of a session, but never again throughout the session. This was important because we needed a task that the participant was likely to complete independently throughout the RSP and ODT condition. Task completion was defined as described above in the experimental conditions. The two tasks the participant completed at moderate levels as compared to the other tasks were chosen to be used during the relevant experimental conditions (i.e., RSP and ODT). Tasks that the participants completed at moderate levels were used because we wanted a task that was likely to occur independently to avoid having to teach the task, but was not likely to occur at such high levels that it interfered with the conditioning procedure (i.e., the participant was attending to the task instead of the relevant pairing).

IOA was calculated for at least 33% of sessions per participant. An agreement for task completion was defined as both observers scoring the occurrence or nonoccurrence of appropriate engagement with a task per interval. IOA was calcuated by dividing the number of agreements by sum of the number of agreements and disagreements and multiplying by 100%. On average, across participants, IOA for mail letters was 100% (range: 97% to 100%), for button press was 81% (range: 52% to 100%), for block-in-button was 98% (range: 90% to 100%), for stack card was 94% (range: 81% to 100%), for card touch was 91% (range: 73% to 100%), and for string bead was 100% (range: 97% to 100%). For Maisy's specific tasks, IOA for drum was 83%, for tambourine was 90%, for piano was 97%, for bells was 100%, for stick was 97%, and for donut was 100%. Because each task was presented three times and IOA was collected for 33% of sessions, IOA was only collected during one session per task. Therefore, there are no average IOA scores or range scores for Maisy.

Identification of sounds. Three, 10 min observations were conducted to identify both target and nontarget speech sounds to be used during the experimental conditions. Observations were conducted in the participant's classroom during free-play periods. All observations were videotaped. Observations were conducted across at least two different days and across at least two different time periods (i.e., morning, midday, and afternoon). No programmed changes occurred during these observations. If leisure items were present in the classroom during the observation, these were available to the participant. Additionally, classroom teachers and peers were present along with the participant and interacted with the participant. A data collector was also present to video tape and stayed in close proximity to the participant, while being as discreet as possible. Two trained observers simultaneously collected data from video using a pencil and paper. Data were collected on the topography of speech sounds, as well as the frequency of each speech sound. Data collectors watched the video repeatedly and discussed the observed speech sounds until they both agreed on the frequency and topography of all observed speech sounds. A topography of speech sound was defined as a simple syllable sound made up of either a vowel sound (e.g., "eee"), a consonant sound ("mmm"), or a combination of the two "duh"). After observers agreed on the topography of a speech sound, they scored the frequency of each speech sound. A pause between speech sounds was not required. Based on the topographies scored, six novel speech sounds were created (i.e., speech sounds not observed to occur during these observations) that were similar to the emitted topographies. For example, if a participant emitted "boo" and "mah," a novel speech sound identified may be "moo." The six novel speech sounds were tested once more for echoic control using the procedures described in the EESA (Esch, 2008). That is, the therapist presented each of the novel speech sounds three times and allowed the participant an opportunity to echo the sound after each presentation. If the participant was unable to echo all of the six novel sounds, then two of the novel speech sounds were quasi-randomly assigned to each of the experimental conditions. Specifically, each sound was assigned a number and then numbers were identified using a random number generator. The sound associated with that number was then assigned to a given condition. However, if sounds were similar to each other (e.g., "duh" and "buh"), these sounds were assigned to different conditions (i.e., SSP, RSP, or ODT) and different trial types (i.e., target or nontarget sound). This was to ensure that any increase observed for a given sound was due to the conditioning procedure and not the topography of the sound. An exception with respect to the ability to echo novel sounds was made for Cynthia, who was the only participant with an emerging echoic repertoire. She was observed to approximate an echoic response for all six novel sounds. We decided to continue to use these novel sounds because she was able to approximate all sounds, suggesting a similar echoic repertoire with respect to these sounds across all conditions. One sound for each condition was assigned as the target sound. The other sound for each condition was assigned as the nontarget sound. IOA was not calculated for observations to identify sounds because both observers verbally agreed on the topographies and frequencies of speech sounds while scoring the videos.

Baseline. A specific baseline condition was conducted for each experimental condition (see Appendix B). The therapist conducted the *SSP baseline* in the yellow room while wearing a yellow shirt. The therapist started a trial by stating, “Look.” If the participant did not orient toward the therapist or relevant stimuli (e.g., task [when applicable], preferred edible) within 3 s of the verbal statement, the therapist prompted the participant to look with as minimally invasive of a prompt as possible without using any further verbal prompts. First, the therapist moved his or her body position to be more directly in front of the participant. If the participant did not orient toward the therapist or relevant stimuli, then the therapist pointed to his or her eyes. If the participant still did not orient toward the therapist or relevant stimuli, then the therapist physically guided the participant’s face to orient toward the therapist. Following an orienting response from the participant, the therapist stated the target or nontarget sound specifically assigned to the SSP experimental condition, three times per trial. A 10 s ITI immediately followed. The therapist continued trials until he or she has conducted 20 trials (i.e., 10 target trials, 10 nontarget trials). The therapist conducted the *RSP baseline* in the purple room while wearing a purple shirt. Additionally, the task identified during the SSTA and assigned to this condition was present. Although the participant was not prompted to engage with the task during baseline, the task was present to ensure that any increase in vocalizations during conditioning was due to the conditioning procedure and not the presence of the task. Pre-session exposure to the task occurred before every session. That is, the therapist verbally asked the child to engage in the task. If the participant did not complete the task, then the therapist modeled the task. If the participant still did not complete the task, then the therapist physically guided the participant to complete the task. During the session, the participant may engage with the task, but the therapist did not prompt the participant to engage with the task. The therapist ensured the task remained within arm's reach of the participant at all times. Trials were identical to the trials during the SSP baseline, except that the therapist stated the target or nontarget sound specifically assigned to the RSP experimental condition, three times per trial. Finally, the therapist conducted the *ODT baseline* in the green room while wearing a green shirt. Again, the task identified during the SSTA and assigned to this condition was present. A different task then the task present in the RSP condition was used to help the participant discriminate between conditions and to ensure that any stimulus control associated with the task was paired with one conditioning procedure. Just as in RSP baseline, the therapist conducted pre-session exposure to the task. During the session, the participant might have engaged with the task, but the therapist did not prompt the participant to engage with the task. The therapist also ensured the task was kept within arm's reach of the participant. Again, trials were identical to the trials during the SSP baseline and the RSP baseline, except that the therapist stated the target or nontarget sound specifically assigned to the ODT experimental condition, three times per trial.

Stimulus-stimulus pairing (SSP). SSP was similar to the SSP baseline condition, except the therapist delivered the preferred edible (i.e., one of the top three preferred edibles identified by the PSPA that the participant chose during a brief multiple-stimulus-without-replacement preference assessment prior to session) simultaneously with the third presentation of the target sound (see Appendix B). The preferred edible was delivered simultaneously with the third presentation of the target sound because these procedures directly matched the enhanced SSP procedures suggested by Esch et al. (2009). The 10 s ITI began after the participant consumed the edible. Following all target trials, the therapist performed a clean-mouth check before beginning the ITI. Specifically, the therapist watched for chewing to cease and then looked into the participant's mouth to visually verify no food was left in the participant's mouth.

The therapist minimized physical contact when performing the clean-mouth check. This ensured that the preferred edible was not available during the nontarget trials, and thus not adventitiously paired with the delivery of the nontarget sound. This also allowed for varied lengths of time between trials, when consumption time was included, such that participants were less likely to be able to predict the delivery of reinforcement. This was also suggested by Esch et al. (2009). However, the set 10 s ITI following consumption ensured that the participant had an equal amount of time to emit vocalizations without food potentially interfering. The therapist conducted nontarget trials identical to trials during SSP baseline. The therapist began the ITI immediately following the nontarget trial. Again, this continued until the therapist conducted 20 trials (i.e., 10 target trials, 10 nontarget trials).

Additionally, an omission contingency was in place to decrease the likelihood that adventitious reinforcement of the target sound would occur. Specifically, if the participant emitted the target sound after the start of the trial, but before the preferred edible was delivered, or if the participant emitted the target sound during the last 5 s of the ITI, there was a 20 s delay, after which the therapist attempted the same trial again. All 20 trials still occurred per session.

Response-stimulus pairing (RSP). RSP was similar to the RSP baseline condition, except the participant completed the task before the sounds were delivered (see Appendix B). Specifically, following the orienting response, the therapist placed the task in front of the participant. If the participant did not complete the task within 3 s of its presentation, the therapist modeled the task for the participant. If the participant still did not complete the task within 3 s of the model prompt, the therapist physically guided the participant to complete the task. After the task was complete (i.e., following the presentation of the task, following the model prompt, or following the physical prompt), the therapist presented either the target or the nontarget sound three times per trial. That is, the participant had to complete the task regardless of whether a target or nontarget trial followed task completion. A modification occurred for Maisy if she did not complete the task within 3 s of its presentation. Maisy was diagnosed with a cortical visual impairment, which made it difficult for her to process stimuli visually. Therefore, instead of a model prompt, the therapist placed Maisy's hand to the task. If Maisy did not complete the task within 3 s following her hand being placed on the task, the therapist delivered a physical prompt like was delivered for the other participants. Following the task response, target and nontarget trials were identical to the target and nontarget trials in SSP. That is, the preferred edible was delivered simultaneously with the third presentation of the target sound during target sound trials. Additionally, the therapist conducted the ITI identically to the SSP condition. Again, this continued until the therapist conducted 20 trials (i.e., 10 target trials, 10 nontarget trials). Finally, as in SSP, the omission contingency was in place to ensure adventitious reinforcement of the target sound did not occur.

Operant discrimination training (ODT). ODT was similar to the ODT baseline condition, except the sound was presented before the participant has an opportunity to engage in the task response (see Appendix B). Specifically, following the orienting response, the therapist stated either the target or nontarget sound three times per trial. This was followed by the presentation of the task in front of the participant. If a target sound was presented and the participant did not complete the task within 3 s of its presentation, the therapist modeled the task for the participant. If the participant still did not complete the task within 3 s of the model prompt, the therapist physically guided the participant to complete the task. After the task was complete (i.e., following the presentation of the task, following the model prompt, or following the physical prompt), the therapist delivered the preferred edible. The same modification that occurred for Maisy during RSP with respect to prompting her to complete the task occurred during ODT. If a nontarget sound was presented, the therapist did not prompt the participant to engage in the task response. The task remained in front of the participant for 10 s or until the participant completed the task response. After 10 s elapsed or the participant completed the task response (whichever comes first), the task was removed and the 10 s ITI began. That is, unlike RSP, the participant did not have to complete the task if the nontarget sound was presented. The participant only had to complete the task if the target sound was presented. This was because ODT relies on the participant discriminating which sound signals reinforcement and which sound does not signal reinforcement, whereas this discrimination is not necessary for RSP. Additionally, for ODT, the therapist conducted the ITI identically to the SSP and RSP conditions. Again, this continued until the therapist conducted 20 trials (i.e., 10 target trials, 10 nontarget trials). Finally, just as in SSP and RSP, the omission contingency was in place to ensure adventitious reinforcement of the target sound did not occur.

Contingency reversal. If one or more experimental condition established a higher rate of target sounds as compared to the other experimental conditions, then a contingency reversal phase occurred using the specific experimental condition or conditions that increased the target sound in order to replicate effects. Specifically, for the condition or conditions that increased target sounds, the therapist placed the contingencies previously associated with the target sound on the nontarget sound and vice versa. For example, if RSP resulted in an increase in the target sound, "Doe," but did not result in an increase in the nontarget sound, "Soo," then during the contingency reversal, the therapist conducted the contingencies associated with the target sound for the sound, "Soo," and not for the sound, "Doe."

Generalization probes. Generalization probes in the session room occurred before and after each phase, and then once approximately every five session blocks. Generalization probes in the classroom occurred before and after each phase, and then once approximately every 15 session blocks. Generalization probes were 10 min in length. During generalization probes in the session room, the therapist delivered noncontingent physical and verbal attention. The therapist did not deliver any demands and did not present the target or nontarget sounds. No programmed consequences occurred. During generalization probes in the classroom, the therapist did not interact with the participant (although teachers and peers might have). No programmed consequences above consequences delivered by the typical teachers in the classroom occurred.

Results

Results for Study 1 are summarized in Table 4 and are graphically depicted in Figures 1-3. In general, we observed three patterns of responding. Figure 1 depicts the first pattern of responding in which at least one conditioning procedure resulted in an increase in the target sound as compared to the nontarget sound. This pattern of responding was observed for three participants, but across two sound sets for one participant. The first graph is for Benny (Sound set 1). When Benny entered this study, her EESA score was zero. During baseline, Benny emitted the target and nontarget sounds at a near zero rate. During the conditioning phase, Benny emitted the target sound at a differentially higher rate as compared to the nontarget sound during both the RSP and SSP condition. At this point, Benny's parent requested that Benny only participate in the RSP condition. Therefore, we only conducted the contingency reversal using the RSP procedures, and did not complete a contingency reversal using the SSP procedures. During the contingency reversal, we observed a differentially higher rate of the target sound as compared to when this same sound was a nontarget sound in the previous conditioning phase.

However, we continued to see differentially higher rates of the nontarget sound, which was previously the target sound. In an attempt to decrease the rate of the nontarget sound, we eliminated the presentation of nontarget trials. Instead, we presented 20 trials of the target sound only. We observed a slight decrease in the nontarget sound, but the rate of the target sound remained the same. Therefore, we presented both target and nontarget trials again. We continued to observe a similar rate of both the target and nontarget sound during the final phase. Across phases, other sounds generally remained the same, with the exception of a decrease during the contingency reversal when nontarget trials were eliminated. Across generalization probes, Benny only consistently emitted the original target sound associated with the RSP condition during the contingency reversal (when this sound was the nontarget sound) in the session room. She emitted this same sound in the classroom during her final generalization probe. We did not conduct the EESA again following the contingency reversal, but waited until after the analysis of Sound set 2 with Benny, at which point Benny's EESA score had increased to three and a half. These data suggest that RSP was an effective conditioning procedure for Benny, and generalization for RSP target sounds occurred, but was delayed. These data may also suggest that SSP was an effective conditioning procedure for Benny, but are limited because we were unable to replicate these results during a contingency reversal.

The second graph is for Benny (Sound set 2). Conditioning procedures were conducted across two sets of sounds with Benny in an attempt to replicate the results we observed with Sound set 1. Two sound sets were only included for Benny. Again, when Benny entered this study, her EESA score was zero. At the request of the parent, we only conducted the RSP condition. During baseline, Benny did not emit the target or nontarget sound. During the first conditioning phase, Benny emitted the target sound at a differentially higher rate as compared to the nontarget sound. When we returned to baseline, Benny stopped emitting the target sound and did not emit the nontarget sound. Finally, when we returned to conditioning (RSP), Benny again emitted the target sound at a differentially higher rate as compared to the nontarget sound. Across phases, we did not see a change in other sounds. Across generalization probes, Benny did not emit the target or nontarget sound from the RSP condition in the session room or in the classroom. Following the final conditioning phase, Benny's EESA score had increased from zero to three and a half. These data suggest that RSP was an effective conditioning procedure for Benny, but generalization did not occur.

The third graph is for Cynthia. When Cynthia entered this study, her EESA score was 17.5. During baseline, Cynthia emitted the target and nontarget sounds at a variable and decreasing rate. During the conditioning phase, Cynthia emitted the target sound at a differentially higher rate as compared to the nontarget sound during both the RSP and SSP condition. To further evaluate the effects of each conditioning procedure, we presented each condition in isolation for several consecutive sessions. During the RSP only condition, Cynthia continued to emit the target sound at a differentially higher rate as compared to the nontarget sound, and at a similar rate as observed during the previous conditioning phase. During the SSP only condition, Cynthia continued to emit the target sound at a differentially higher rate as compared to the nontarget sound, and at a similar rate as observed during the previous conditioning phase. During the ODT only condition, Cynthia began to emit the target sound at a differentially higher rate as compared to the nontarget sound. This was not observed during the previous conditioning phase. Unfortunately, during the contingency reversal, we did not replicate the results observed during the conditioning phase. That is, Cynthia did not emit the target sound at a differentially higher rate as compared to the nontarget sound. In fact, she rarely emitted the target or the nontarget sound across all three conditioning procedures. Across phases, we did not see a change in other sounds. Across generalization probes, Cynthia consistently emitted the target sounds associated with the RSP and SSP conditions in the session room. She also emitted the target sounds associated with the RSP and SSP conditions in the classroom during the last generalization probe. Following the final contingency reversal phase, Cynthia's EESA score increased from 17.5 to 18.5. These data suggest that RSP, SSP, and ODT were effective conditioning procedures for Cynthia initially, and generalization occurred for RSP and SSP target sounds, but not for ODT target sounds. These results are limited because we were unable to replicate conditioning effects during the contingency reversal phase. Therefore, it is unclear if extraneous variables are responsible for the increase in the target sound during the initial conditioning phases.

The fourth graph is for Allan. During baseline, Allan emitted the target sound in the RSP condition, but at a near zero rate. During the conditioning phase, Allan emitted the target sound in the ODT condition at a differentially higher rate than the nontarget sound. Although the nontarget sound in the ODT condition did occur at the beginning of the phase, this stopped by the end of the phase. To further evaluate the effects of the ODT condition, we conducted ODT in isolation for several consecutive sessions. We continued to see differentially higher rates of the target sound as compared to the nontarget sound, and the target sound occurred more consistently. Therefore, we conducted a contingency reversal phase using the ODT procedures. During this phase, we replicated the results of the previous conditioning phase. That is, we saw differentially higher rates of the target sound (previous nontarget sound) as compared to the nontarget sound (previous target sound). Across phases, other sounds generally remained the same. Across generalization probes, Allan emitted the target sounds associated with the RSP condition and the ODT condition both in the session room and the classroom. He also emitted the nontarget sound associated with the SSP condition during his last two session room generalization probes. These data suggest that ODT was an effective conditioning procedure for Allan, and generalization for ODT target sounds occurred, but generalization also occurred for the RSP target sound and the SSP nontarget sound.

Figure 2 depicts the second pattern of responding in which none of the conditioning procedures resulted in an increase in the target sound as compared to the nontarget sound. The first graph is for Mason. During baseline, we observed a near zero rate of responding across all sounds and all conditions. During the conditioning phase, we continued to observe a near zero rate of responding across all sounds and all conditions. Because the majority of the literature describes SSP, we evaluated SSP in isolation for several consecutive sessions in an attempt to gather more information specifically on this procedure. First, we returned to baseline and again Mason emitted the target and nontarget sounds at a near zero rate across all conditions. Then, we conducted baseline for SSP only. Mason did not emit the target or nontarget sound. Finally, we conducted conditioning for SSP only. Mason did not emit the target sound and emitted the nontarget sound at a near zero rate. Across phases, other sounds generally remained the same. During generalization probes, Mason emitted the nontarget sound associated with the RSP condition once. These data suggest that none of the conditioning procedures were effective for Mason, and generalization did not occur.

The second graph is for Antonio. During baseline, we observed a near zero rate of responding across all sounds and all conditions. During the conditioning phase, we continued to observe a near zero rate of responding across all sounds and all conditions. Like with Mason, we evaluated SSP in isolation for several consecutive sessions in an attempt to gather more information specifically on this procedure. Interestingly, during baseline for SSP only, we observed a differentially higher rate of the target sound as compared to the nontarget sound, but this decreased to zero across the phase. When we conducted conditioning for SSP only, the target sound did not occur. Across phases, other sounds generally remained the same. During generalization probes, Antonio emitted the target sound associated with SSP during one session room generalization probe and during one classroom generalization probe. These data suggest that none of the conditioning procedures were effective for Antonio, and generalization did not occur.

The third graph is for Maisy. During baseline, we observed a near zero rate of responding across all sounds and all conditions. During the conditioning phase, we continued to observe a near zero rate of responding across all sounds and all conditions. Across phases, other sounds generally remained the same. During generalization probes, Maisy did not consistently emit any of the sounds. These data suggest that none of the conditioning procedures were effective for Maisy, and generalization did not occur.

Figure 3 depicts the third and final pattern of responding in which none of the conditioning procedures resulted in an increase in the target sound as compared to the nontarget sound during the ITI, which is when an increase should have been observed if a conditioning procedure successfully resulted in conditioning a sound as a reinforcer. However, when we analyzed the rate of the target and nontarget sounds during the trial, we observed an increase during the trial – suggesting this participant's echoic repertoire increased throughout the evaluation. This pattern of responding was only observed for one participant – Octavia. The top graph depicts Octavia's responding during the ITI. During baseline, Octavia emitted target and nontarget sounds at a variable rate across all conditions. During the conditioning phase, Octavia continued to emit the target and nontarget sounds at a variable rate across all conditions. Across phases, other sounds generally remained the same. Across generalization probes, Octavia emitted the target sound associated with ODT during one session room generalization probe and the target sound associated with SSP during one classroom generalization probe. These data would suggest that none of the conditioning procedures were effective for Octavia, but generalization of the target sound for ODT and SSP occurred (although it was limited). However, next we analyzed the rate of Olivia's responding during the trial, as opposed to during the ITI. The bottom graph depicts Octavia's responding during the trial. During baseline, Octavia emitted the target and nontarget sounds at a variable rate across all conditions. During the conditioning phase, Octavia began to emit both the target and nontarget sound across all three conditions, except the nontarget sound associated with the SSP condition. These data suggest that exposure to the conditioning procedures may have increased Octavia's echoic repertoire. However, these data are limited because we did not replicate the results so it is unclear if the conditioning procedures increased Octavia's echoic repertoire or if acquired an echoic repertoire due to an extraneous variable.

Study 2

Purpose

Participants for whom all three conditioning procedures were unsuccessful at increasing the frequency of novel vocalizations participated in Study 2 (i.e., Maisy, Mason, and Antonio). The purpose of Study 2 was to evaluate potential variables that may hinder the effectiveness of conditioning procedures (SSP, RSP, and ODT). Petursdottir and Lepper (2015) offered several potential reasons for treatment failures. First, the inability to establish vocalizations because of speech production problems (i.e., the child is unable to articulate the correct sounds). Second, the primary reinforcer used in the conditioning procedure may not be sufficiently powerful. Third, the child may lack the required attending repertoire. Therefore, we attempted to evaluate the potential existence of articulation problems, insufficient reinforcer potency, and limited attending skills with the three participants for whom no conditioning procedure was effective for increasing target vocalization in Study 1.

To test for articulation, a post-hoc analysis of 50% of sessions per condition from Study 1 for a given participant was conducted. The frequency of approximations to the target and nontarget sounds was recorded. These data were graphed in a multielement design identical to Study 1. Additionally, to further test articulation the methodology of Study 1 was replicated, except known (instead of novel) sounds were used such that it was clear that the participant could fully emit a target and nontarget sound used in the conditioning procedures. To test for reinforcer efficacy, a reinforcer assessment was conducted using a reversal design. Finally, to test for attending behavior, a post-hoc analysis of 50% of sessions per condition from Study 1 for a given participant was conducted. The total prompts necessary to obtain an orienting response as well as the latency to the orienting response was recorded. These data were graphed in a multielement design identical to Study 1.

Phase 1: Articulation

Response measurement and reliability. In Study 1, the occurrence of a target or nontarget sound were defined as the participant emitting a sound that exactly matched the target or nontarget sound. It is possible that some participants emitted approximations to the target and nontarget sounds, but these did not meet the definition for data collection. Therefore, a post-hoc analysis of 50% of sessions per condition from Study 1 for approximations was conducted. If a participant approximated target sounds, then we concluded that articulation may be a variable that decreased the effectiveness of the conditioning procedures. Data were collected on the occurrence of a given behavior within a 5 s interval. Specifically, data were collected on the occurrence of the target sound, the nontarget sound, an approximation to the target sound, an approximation to the nontarget sound, and any other sound emitted by the participant. The occurrence of the target sound and nontarget sound were defined identically to Study 1. An approximation to a target or nontarget sound was defined as any speech sound that contained either a vowel or a consonant (not both) that required the same proximate placement of the tongue or lip as the target or nontarget sound, and the consonant or vowel were in the same order as the target or nontarget sound. Finally, any other sound was defined identically to Study 1.

To help observers quickly identify if a sound met the definition of an approximation, all possible approximations for a target or nontarget sound for a given participant were identified using a chart and information provided by Peña-Brooks and Hedge in the textbook, Assessment and Treatment of Articulation and Phonological Disorders in Children (2007). The chart identified all consonants that were similar based on proximate placement of the tongue (see Appendix C). Peña-Brooks and Hedge (2007) did not provide a similar chart for vowels. However, these authors provided information such that we could make our own chart pertaining to vowels that were similar based on proximate placement of the tongue (see Appendix C). Observers referenced the list of all possible approximations while scoring videos.

Trained observers collected data using ABC Data Pro on iPods. Data collection occurred solely from video and included 50% of sessions per condition per participant. Sessions were from Study 1 and were selected using a random number generator. To ensure an equal spread of sessions across conditions, sessions were split into blocks of five. If the total number of sessions was not divisible by five, the extra or missing sessions were placed in the final session block.

Then, an equal number of sessions was selected from each session block until 50% of the total number of sessions for a given condition were chosen. If the number of sessions could not be evenly selected across session blocks, the extra session was randomly chosen from any session block. For example, if SSP in baseline consisted of 14 sessions, the session blocks identified were sessions 1-5, 6-10, and 11-14. To score 50% of sessions for this condition, seven total sessions were scored. To ensure an equal number of sessions across session blocks, two sessions from each block would be scored, with one extra session identified from any session block. Using a random number generator, the following sessions were chosen: sessions 2, 3, 6, 8, 10, 12, and 14. A second independent observer collected data independently during at least 33% of the 50% of sessions scored for articulation to assess IOA. Partial interval reliability was calculated by dividing the number of agreement intervals by the total number of intervals and multiply by 100%. On average, across participants, IOA for the target sound was 100% (range: 98% to 100%), for the nontarget sound was 100% (range: 99% to 100%), for approximations to the target sound was 100% (range: 97% to 100%), for approximations to the nontarget sound was 98% (range: 81% to 100%), and for other sounds was 89% (range: 80% to 96%).

Procedural fidelity. Procedural fidelity data were not collected while scoring approximations because these data were collected during Study 1.

Articulation test with known sounds. It is possible that participants may not approximate sounds, but articulation is still a variable decreasing the efficacy of the conditioning procedures because a participant may not have the ability to even approximate a given sound. Therefore, a further test of articulation was conducted by replicating the methodology from Study 1, except known sounds were used to eliminate any articulation problem. The purpose of this test was to ensure that articulation was not a variable affecting the efficacy of SSP, RSP, or

ODT. To do so, sounds were identified that the participant already demonstrated the ability to emit. If the conditioning procedures did not increase target sounds when known sounds were used, then we concluded that articulation was not a variable that decreased the effectiveness of the conditioning procedures because conditioning procedures failed to increase vocalizations, even when these vocalizations were sounds the participant could articulate.

Response measurement and reliability. Data were collected on all dependent variables (i.e., target sounds, nontarget sounds, and other sounds) identically to Study 1. Additionally, data were collected on task responses identically to Study 1. All operational definitions remained the same, as well as data collection methods (i.e., BDataPro). Finally, reliability was calculated identically to Study 1. Across known sounds during Study 1 (described below), baseline, conditioning, and contingency reversals, the average IOA across participants for the target sound was 97% (range: 89% to 100%), for the nontarget sound was 96% (range: 87% to 100%), for other sounds was 80% (range: 62% to 86%), and for task response was 94% (range: 88% to 100%). Across classroom and session room generalization probes, the average IOA across participants for the SSP target sound was 97% (range: 89% to 100%), for SSP nontarget sound was 99% (range: 95% to 100%), for RSP target sound was 97% (range: 91% to 100%), for RSP nontarget sound was 97% (range: 93% to 100%), for ODT target sound was 100% (range: 98% to 100%), for ODT nontarget sound was 98% (range 95% to 100%), and for other sounds was 88% (range: 75% to 100%).

Procedural fidelity. Data on therapist behavior (i.e., delivery of the target and nontarget sound, delivery of the preferred item, length of ITI, therapist prompts) were collected identically to Study 1. Additionally, the level of procedural fidelity was calculated identically to Study 1. On average, across participants, the percentage of sessions that included the correct number of target and nontarget sessions was 93% (range: 86% to 100%), the percentage of trials that were implemented correctly was 90% (range: 85% to 94%), and the percentage of trials during which the omission contingency was implemented correctly was 80% (range: 67% to 93%).

Pre-assessments. The preferred items previously identified by the PSPA and the task items previously identified by the SSTA during Study 1 were used. Therefore, these pre-assessments were not conducted again. Known sounds (i.e., sounds that the participant was observed to emit in his or her classroom and also during previous experimental sessions in Study 1) were identified similarly to Study 1.

Identification of sounds. Known sounds were identified identically to Study 1. That is, three, 10 min observations were conducted in the participant's classroom during free-play periods. However, in addition to the three, 10 min observations in the participant's classroom, three previous sessions from Study 1 were also scored with respect to the topography of speech sounds emitted by the participant. One session from each condition (SSP, RSP, and ODT) from Study 1 were identified to score by placing all numbers associated with a given condition into a random number generator and using the first number generated. However, at least one session was a baseline session and at least one session was a conditioning session. Thus, if more than two sessions were identified by the random number generator for baseline or conditioning, new numbers were generated until an appropriate session was identified.

Across both the classroom observations and the previous sessions, two trained observers collected data simultaneously using a pencil and paper. Like in Study 1, data were collected on the topography of speech sounds, as well as the frequency of each speech sound. The operational definitions remained the same. Also like in Study 1, data collectors watched the videos repeatedly and discussed the observed speech sounds until they both agreed on the topography and frequency of each speech sound.

Based on the topographies scored, six known speech sounds that occurred at least once across the classroom observations and the previous sessions were selected. This ensured sounds that the participant emitted in both settings were selected. The six known speech sounds were tested once more for echoic control using the procedures described in the EESA (Esch, 2008). If the participant was unable to echo each of the six known sounds, then two of the known sounds were randomly assigned to each of the experimental conditions. Like in Study 1, one sound for each condition was assigned as the target sound and the other sound for each condition was assigned as the nontarget sound. Similar to Study 1, IOA was not calculated for observations to identify sounds because both observers verbally agreed on the topographies and frequencies of speech sounds while scoring the videos.

Known sounds occurrence during Study 1. To get a baseline rate of the six known sounds identified, 33% of sessions per condition from Study 1 were scored with respect to the known target and nontarget sounds, and other sounds. Sessions were identified using a random number generator. Sessions were scored identically to Study 1.

Baseline. Baseline was conducted identically to Study 1, except known sounds were used instead of novel sounds. The same tasks as in Study 1 were paired with the same conditions as in Study 1. Additionally, the same session rooms were used as in Study 1 with the room colors correlated with the same conditions as in Study 1.

SSP, RSP, and ODT. SSP, RSP, and ODT were conducted identically to Study 1, except known sounds were used instead of novel sounds. As in Study 1, a choice of the top three preferred edibles from the PSPA were offered to the participant before the start of each session.

The item that the participant chose was paired with the known sound during the session. Again, the tasks and session rooms from Study 1 were paired with the same conditions.

Contingency reversal. As in Study 1, if one or more experimental condition had established a higher rate of target sounds as compared to nontarget sounds, then a contingency reversal phase would have been conducted using the specific experimental conditions that increased target sounds in order to replicate effects. However, for the two participants for whom we completed the articulation test with known sounds (i.e., Mason and Antonio), none of the experimental conditions established a higher rate of target sounds as compared to nontarget sounds.

Generalization probes. As in Study 1, generalization probes were periodically conducted in the session room and in the classroom. Generalization probes were conducted on the same schedule as in Study 1 (i.e., before and after each phase, and then in a session room approximately once every five session blocks and in a classroom approximately once every 15 session blocks). Generalization probes were 10 min in length.

Phase 2: Reinforcer Efficacy

Response measurement and reliability. The test for reinforcer efficacy was only conducted with Maisy because Mason and Antonio left the study before we were able to test for reinforcer efficacy. During Study 1, the preferred items were never delivered contingent on the participant's behavior. Instead, they were paired with the therapist's behavior, as indicated by a given conditioning procedure. Therefore, the purpose of the test for reinforcer efficacy was to evaluate whether the preferred items functioned as reinforcers when delivered contingent on a participant's behavior. The primary dependent variable was the frequency of responses (i.e., target sound, a nonverbal behavior – specifically hitting a tambourine) or the percentage of intervals in which a response occurs (i.e., for all sounds). Responses were graphed as a frequency for target sound and nonverbal behavior or as a percentage of intervals per minute for all sounds, with minutes on the x-axis for all three behaviors. A target sound was defined identically to Study 1. All sounds were defined identically to other sounds in Study 1. The nonverbal behavior was defined identically to Study 1. Specifically, for Maisy, the nonverbal behavior was hitting the tambourine. Trained observers collected data using BDataPro. Data collection occurred in-vivo or from video. A second independent observer collected data independently during all sessions. Proportional reliability was calculated for the frequency of responses by dividing the sum of the total agreement intervals and the proportion of agreement per interval by the total number of intervals and multiplying by 100%. Partial interval reliability was calculated by dividing the number of agreement intervals by the total number of intervals and multiplying by 100%. On average, across baseline, FR1, and EXT, IOA for the target sound was 99%, for all sounds was 79%, and for the nonverbal behavior (i.e., hitting the tambourine) was 79%. Because Maisy was the only participant to participate in this phase, there were no ranges for IOA scores.

Procedural fidelity. During the reinforcer assessments, data were collected on the therapist behavior with respect to the delivery of the preferred edible. The delivery of the preferred edible was defined identically to Study 1. The level of procedural fidelity was calculated by identifying the total number of preferred edibles delivered and the total number of target behaviors emitted by the participant and dividing the smaller by the larger. On average, across all three reinforcer assessments, procedural fidelity was 96% (range: 73% to 100%).

Experimental conditions. Three reinforcer assessments were conducted using three different target behaviors. The target behavior for the first reinforcer assessment was the target or nontarget sound emitted most frequently by Maisy during Study 1. This sound was chosen because it was most likely to occur during a reinforcer assessment such that Maisy was likely to contact reinforcement, even without pre-session exposure to the contingencies (which is important because it is not possible to physically prompt a vocal response). Because a reinforcer effect was not observed (i.e., Maisy was not more likely to emit the given sound when the preferred item was delivered contingently), the second reinforcer assessment was conducted. It is possible that the preferred items were not strong enough reinforcers to increase a given sound, but could still function as reinforcers for other vocal behavior. Therefore, the target behavior for the second reinforcer assessment was all sounds. This was chosen as a target behavior because it increased the likelihood that Maisy would contact reinforcement without pre-session exposure to the contingencies (which is important because it is not possible to physically prompt a vocal response). Because a reinforcer effect was not observed, we conducted the third reinforcer assessment. It is possible that the preferred items were not strong enough reinforcers to increase any vocal response, but could still function as reinforcers. Therefore, the target behavior for the third reinforcer assessment was a nonverbal behavior. Specifically, it was the task that Maisy completed most frequently across sessions in Study 1. This task was chosen because it was most likely to occur during a reinforcer assessment such that Maisy was likely to contact reinforcement, even without pre-session exposure to the contingencies. We did not conduct pre-session exposure during this reinforcer assessment, even though it was possible to physically prompt this task, because we wanted to keep the absence of pre-session exposure consistent across reinforcer assessments. If a reinforcer effect was not observed during all three reinforcer assessments, then we concluded that reinforcer efficacy was a variable that may decrease the effectiveness of the conditioning procedures.

Identically to Study 1 for Maisy, one of the top three preferred items identified during the PSPA was used during each reinforcer assessment. One of the top three preferred items was quasi-randomly chosen such that each item was used once across the three reinforcer assessments. This was to ensure that all three preferred items were tested for reinforcer efficacy. If the delivery of a preferred item did not result in an increase in a target behavior, then this preferred item may not be a reinforcer, or at least may not be a potent enough reinforcer to increase the given target behavior. Each reinforcer assessment consisted of three phases: baseline, fixed-ratio 1 (FR1), and extinction (EXT). Each phase lasted 10 minutes and was conducted consecutively. That is, there was no break in time between phases, so one session was 30 min in length.

Baseline. During baseline, the therapist was present and the preferred item was in view. The preferred item was in view in baseline to control for its presence during the FR1 condition. It is possible that simply the presence of a preferred item could increase a behavior and we wanted to demonstrate that any increase in behavior was due to the delivery of the preferred item and not merely its presence. As in Study 1, the therapist blocked any attempt to access the preferred item. During the first reinforcer assessment, the therapist stated, "Say (target sound)" every 10 s. During the second and third reinforcer assessment, the therapist ignored Maisy. If Maisy engaged in the target behavior during any of the reinforcer assessments, no programmed consequences occurred.

FR1. During FR1, the therapist and preferred item continued to be present and in view. The therapist continued to block any attempt to access the preferred item. During the first reinforcer assessment, the therapist continued to state, "Say (target sound)" every 10 s. During the second and third reinforcer assessment, the therapist continued to ignore Maisy. However, if

Maisy engaged in the target behavior during any of the reinforcer assessments, the therapist immediately delivered the preferred item.

Extinction (EXT). EXT was identical to baseline. That is, the therapist and the preferred item were still present. During the first reinforcer assessment, the therapist continued to state, "Say (target sound)" every 10 s. During the second and third reinforcer assessment, the therapist continued to ignore Maisy. Across all three assessments, there were no programmed consequences for the Maisy's behavior.

Phase 3: Attending Behavior

Response measurement and reliability. Petursdottir and Lepper (2015) suggested that an attending repertoire may be important for conditioning procedures to successfully increase vocalizations because participants must attend to the relevant stimuli (i.e., sound presentation, reinforcer delivery). Therefore, the participant's attending repertoire was measured by comparing the number of prompts necessary to obtain an orienting response (as defined in Study 1) as well as the latency to the first sound or task presentation. These variables were chosen as a measure of attending because more prompts and a longer latency to an orienting response should indicate that a participant is less likely to independently attend on command, suggesting attending skills may be weak. Additionally, fewer prompts and a shorter latency to an orienting response means there is less time for confounding variables to occur, making it more likely that the participant attends to the relevant stimuli (i.e., the sound and reinforcer delivery). If a participant required a similar level of prompts or an increased level of prompts during conditioning procedures as compared to baseline or if a similar latency or an increased latency to the first sound or task presentation was observed during conditioning procedures as compared to baseline, then we concluded that attending may be a variable that decreased the efficacy of the conditioning procedures.

Data were collected during a post-hoc analysis of 50% of sessions per condition from Study 1. The primary dependent variables were the total number of prompts necessary to achieve an attending response (i.e., an orienting response as defined in Study 1) from the participant following the statement, “look” for a given session and the average latency to the first sound or task presentation for a given session. A therapist prompt was defined as the therapist moving his or her body position to be more directly in front of the participant, pointing to his or her eyes, or physically guiding the participant’s face to orient toward the therapist. This was scored as a frequency. Latency to the first sound or task presentation was defined as the time between the therapist stating, “look” and either the delivery of the first sound (during SSP and ODT) or the delivery of the task (during RSP). This was scored as a duration. Trained observers collected data using BDataPro. Data collection occurred solely from video and included the same 50% of sessions per condition per participant as were scored during Phase 1 for articulation. A second independent observer collected data independently during at least 33% of the 50% of sessions scored for attending behavior to assess IOA. Proportional reliability was calculated for total prompts by dividing the sum of the total agreement intervals and the proportion of agreement per interval by the total number of intervals and multiplying by 100%. Total reliability per trial was calculated for latency. Latency agreement was defined as the same latency for a given trial, plus or minus three seconds. Total reliability was calculated by dividing the sum of the total agreement trials by the total number trials and multiplying by 100%. On average, across participants, IOA for therapist prompts was 93% (range: 87% to 99%) and for latency was 97% (range: 75% to 100%).

Procedural fidelity. Procedural fidelity data were not collected while scoring attending behavior because these data were collected during Study 1.

Results

Results for Study 2 are summarized in Table 5 and depicted in Figures 4-10. Phase 1: Articulation data are depicted in Figures 4-6. Phase 2: Reinforcer Efficacy data are depicted in Figure 7. Finally, Phase 3: Attending data are depicted in Figures 8-10. Figure 4 depicts data for Mason in Phase 1: Articulation. The top graph depicts the percentage of intervals during which Mason emitted the target sound, nontarget sound, an approximation to the target or nontarget sound, or any other sound. During baseline, Mason did not emit the target or nontarget sounds, but did emit approximations to target and nontarget sounds at a near zero rate. During conditioning, Mason still emitted approximations to some nontarget sounds, but he also consistently emitted approximations to the target sound associated with the ODT condition at an increasing rate. When we returned to baseline, Mason continued to emit approximations to some nontarget sounds, but he stopped emitting approximations to the ODT target sound. Unfortunately, we only conducted the SSP condition during the second conditioning phase. Therefore, there was no opportunity to evaluate whether approximations to the ODT target sound would occur again. Across phases, other sounds generally remained the same. These data suggest that potentially ODT would have increased target sounds, but articulation may have been a barrier. These data are limited because there was no opportunity to replicate an increase in approximations to the ODT target sound in the second conditioning phase.

The bottom graph depicts the level of target, nontarget, and other sounds during the ITI when conditioning procedures were used with known sounds (instead of novel). The first two phases depict known sounds emitted during 33% of the sessions from Study 1. During baseline, Mason emitted a variable rate of target and nontarget sounds across conditions. This pattern of responding was also observed during conditioning. The second two phases depict known sounds emitted during Study 2. During baseline, Mason continued to emit a variable rate of target and nontarget sounds across conditions. This pattern of responding was also observed during conditioning. Across phases, other sounds generally remained the same. Across generalization probes, we observed a variable rate of target and nontarget sounds across conditions. These data suggest that articulation may not be the only barrier for Mason.

Figure 5 depicts data for Antonio in Phase 1: Articulation. The top graph depicts the percentage of intervals during which Antonio emitted the target sound, nontarget sound, an approximation to the target or nontarget sound, or any other sound. During baseline, Antonio rarely emitted the target or nontarget sounds, but did emit approximations to target and nontarget sounds. During conditioning, Antonio's responding decreased and he no longer emitted approximations to target or nontarget sounds. When conducted several consecutive sessions of SSP in isolation, Antonio rarely emitted target sounds, nontarget sounds, or approximations to these sounds. Across phases, other sounds generally remained the same. These data suggest that articulation may not be a barrier for Antonio. These data are limited because it is unclear whether articulation is not a barrier or whether articulation is such a barrier that even approximations do not occur.

The bottom graph depicts the level of target, nontarget, and other sounds during the ITI when conditioning procedures were used with known sounds (instead of novel). The first two phases depict known sounds emitted during 33% of the sessions from Study 1. During baseline, Antonio emitted a variable rate of target and nontarget sounds across conditions. This pattern of responding was also observed during conditioning. The second two phases depict known sounds emitted during Study 2. During baseline, Antonio continued to emit a variable rate of target and nontarget sounds across conditions. This pattern of responding was also observed during conditioning. Across phases, other sounds generally remained the same. Across generalization probes, we observed a variable rate of target and nontarget sounds across conditions.

Figure 6 depicts data for Maisy in Phase 1: Articulation. The top graph depicts the percentage of intervals during which Maisy emitted the target sound, nontarget sound, an approximation to the target or nontarget sound, or any other sound. During baseline, Maisy did not emit the target or nontarget sounds or approximations to target and nontarget sounds. During conditioning, Maisy's responding was similar. She rarely emitting the target or nontarget sounds or approximations to the target or nontarget sounds. Across phases, other sounds generally remained the same. These data suggest that articulation may not be a barrier for Maisy. These data are limited because it is unclear whether articulation is not a barrier or whether articulation is such a barrier that even approximations do not occur.

The bottom graph depicts the level of target, nontarget, and other sounds during the ITI when conditioning procedures were used with known sounds (instead of novel). The first two phases depict known sounds emitted during 33% of the sessions from Study 1. During baseline, Maisy emitted a variable rate of target and nontarget sounds across conditions. This pattern of responding was also observed during conditioning. The second two phases depict known sounds emitted during Study 2. During baseline, Maisy continued to emit a variable rate of target and nontarget sounds across conditions. This pattern of responding was also observed during conditioning. Across phases, other sounds generally remained the same. Across generalization probes, we observed a variable rate of target and nontarget sounds across conditions. Again, these data suggest that articulation may not be a barrier for Maisy.

Figure 7 depicts data for Maisy in Phase 2: Reinforcer Efficacy. The top graph depicts the reinforcer assessment in which the target behavior was the target sound she emitted most frequently in Study 1. During baseline, Maisy did not emit the target sound. During reinforcement, Maisy emitted the target sound once, but then did not emit the target sound again. Finally, during EXT Maisy did not emit the target sound. The middle graph depicts the reinforcer assessment in which the target behavior was all sounds. During baseline, Maisy emitted sounds at a variable level. During reinforcement, Maisy continued to emit sounds at a similar variable level. Finally, during EXT Maisy continued to emit sounds at a similar variable level. The bottom graph depicts the reinforcer assessment in which the target behaviors was the nonverbal behavior (or task) that Maisy engaged with most frequently during Study 1. During baseline, Maisy engaged in the nonverbal behavior at a variable level. During reinforcement, Maisy engaged in the nonverbal behavior, but at a decreased level as compared to baseline. During EXT, Maisy again engaged in the nonverbal behavior at a level similar to baseline. Together, these data suggest that reinforcer efficacy may be a barrier for Maisy. In fact, the delivery of the preferred item seemed to decrease responding.

Figure 8 depicts data for Mason in Phase 3: Attending. During baseline, we observed an increase in latency to the first sound presentation or task presentation as well as an increase in the total number of prompts needed to facilitate attending. During conditioning, we observed a decrease in both the latency and total prompts. When we returned to baseline, we again observed an increase in latency and total prompts. However, when we returned to conditioning, latency and total prompts remained the same. These data suggest that attending may not be a barrier for Mason because attending improved during the conditioning phase (i.e., the latency was shorter and there were fewer total prompts). These data are limited because the decrease in latency and total prompts was not replicated during the second conditioning phase.

Figure 9 depicts data for Antonio in Phase 3: Attending. During baseline, we observed a steady duration in latency to the first sound presentation or task presentation as well as a steady level in the total number of prompts needed to facilitate attending. During conditioning, we observed a similar level in both latency and total prompts. When we conducted baseline for SSP only, we again observed a similar level of latency and total prompts. Finally, when we returned to conditioning for SSP only, we again observed a similar level of latency and total prompts. These data suggest that attending may be a barrier for Antonio because attending did not improve during the conditioning phase (i.e., the latency and total prompts remained the same).

Figure 10 depicts data for Maisy in Phase 3: Attending. During baseline, we observed a steady duration of latency to the first sound presentation or task presentation as well as a steady level in the total number of prompts needed to facilitate attending. During conditioning, we observed a decrease in total prompts, but latency remained about the same. These data suggest that attending may not be a barrier for Maisy because attending improved during the conditioning phase (i.e., there were fewer total prompts).

Discussion

The results of the current study were idiosyncratic. That is, we did not consistently see that any one procedure was more likely to increase vocalizations, and the barriers that appeared likely to influence the efficacy of conditioning procedures were different for different children. During Study 1, at least one conditioning procedure increased vocalizations during the ITI for three of the seven participants (both sounds sets for Benny, Cynthia, and Allan). SSP resulted in an increase in vocalizations for two participants (Cynthia and Benny’s sound set 1). RSP resulted in an increase in vocalizations for two participants (both sound sets for Benny and Cynthia). ODT resulted in an increased in vocalizations for two participants (Cynthia and Allan). For one participant (Octavia), an increase in vocalizations did not occur during the ITI. However, she emitted both the target and nontarget sounds across all three conditions during the trial, suggesting her echoic repertoire increased. For the remaining three participants (Maisy, Mason, and Antonio), none of the conditioning procedures resulted in an increase in vocalizations.

With respect to generalization, we did see at least some generalization to the session room or the classroom with four of the participants (Cynthia, Benny’s sound set 1, Allan, and Octavia). It is interesting to note that these are the four participants for whom we saw an increase during at least one of the conditioning procedures, albeit during the trial time for Octavia. Additionally, we observed an increase in the EESA score following conditioning procedures for four participants (Benny’s sound set 2, Cynthia, Allan, and Octavia; See Table 2 for post-conditioning EESA scores). Again, this was with the four participants for whom we saw in increase during at least one of the conditioning procedures.

With respect to Study 2, at least one barrier was identified for each participant for whom none of the conditioning procedures resulted in an increase in vocalizations. Specifically, reinforcer efficacy was identified as a barrier for Maisy, articulation was identified as a barrier for Mason, and attending was identified as a barrier for Antonio.

Taken together, the results of Study 1 and 2 may suggest that a given conditioning procedure may be more effective for a particular child, and sometimes none of the conditioning procedures will be effective. However, when a conditioning procedure is effective, it seems likely that not only will these effects generalize to other environments, but also that the child’s echoic repertoire will increase (at least as indicated by an EESA score). Additional treatment comparison studies seem warranted to provide more information regarding relative effectiveness of SSP, RSP, and ODT procedures. Until these data are available, our multielement treatment comparison could be used to in practical application as an initial assessment to identify an effective conditioning procedure for a given child. However, one potential limitation is multiple treatment interference as a result of exposure to rapidly alternating experimental conditions such that individual treatment effects are masked. For example, during the first conditioning phase for Cynthia a multielement design was used and ODT appeared ineffective. When we conducted consecutive sessions of a single condition, ODT was effective. Although a multielement design allows for a quick analysis of all three conditioning procedures, interaction effects across conditioning procedures may result in an inaccurate reflection of the effectiveness of conditioning procedures for a given participant (Hains & Baer, 1989). Researchers might consider formally evaluating the extent to which multielement designs are associated with multiple treatment interference within the context of conditioning child vocalizations.

Another important area for future research is to identify prerequisite skills that may be necessary for one or more of these conditioning procedures to be effective. Table 3 displays the skill repertoire of participants before participating in the current study. Unfortunately, we did not systematically conduct skill assessments with all participants prior to the start of the study. However, these data were available for the four participants for whom these assessments were conducted as part of their clinical treatment. Although these data are limited, it does suggest that imitation and vocal skills were higher for Benny (for whom vocalizations increased) as compared to Mason, Antonio, and Maisy (for whom vocalizations did not increase). Future research may consider systematically evaluating such skills as potential prerequisite skills that may be necessary for one or more of these conditioning procedures to be effective.

Study 2 of the current study also attempted to initially identify prerequisite skills that may be important when using conditioning procedures by evaluating articulation and attending skills as potential barriers, but there were several limitations. First, we only completed these analyses for three participants. Articulation was identified as a barrier for one participant and attending was identified as a barrier for the other participant. Therefore, it would be helpful to complete these analyses with a larger sample of participants. Second, we conceptualized an articulation barrier as being present when a participant approximated the target sound (as we saw with Mason). Such responding would suggest that potentially a given conditioning procedure did increase vocalizations, but our measurement system was not sensitive enough to capture this increase. However, zero approximations (as we saw with Antonio) could also be conceptualized as an articulation barrier because articulation might be such a large barrier that a participant may not even be able to emit an approximation. Third, we conceptualized an attending barrier as being present if the therapist presented the first sound or the task faster or used fewer prompts during the conditioning phase as compared to baseline. However, this assumes that the therapist correctly presented the sound or the task after obtaining an attending response (i.e., the participant orients his or her eyes toward the therapist or materials) and assumes that the therapist correctly prompted the participant until obtaining an attending response. We were unable to collect data on attending responses directly because observers were behind an observation window and unable to see the participant's eyes on a consistent basis. Therefore, it is unclear whether data on the therapist's behavior is an accurate reflection of the participant's ability to attend.

Researchers should consider additional methods for evaluating barriers such as articulation and attending. For example, researchers may consider testing conditioning procedures, then teaching skills that may be pre-requisites, and then testing conditioning procedures again to see whether these procedures are more effective after pre-requisite skills have been obtained. Additionally, researchers may consider evaluating other potential barriers. For example, auditory discrimination may be an important prerequisite skill for conditioning procedures. Benasich and Tallal (2002) found that an infant's ability to discriminate rapid auditory cues was a good predictor of later language outcome. Thus, researchers might consider teaching auditory discrimination before conducting conditioning procedures to see whether conditioning procedures are more effective after auditory discrimination training occurs. Finally, researchers should also consider evaluating procedural differences (e.g., number of sound presentations) and characteristic differences (e.g., vocal repertoire) as potential reasons a given conditioning procedure might fail to be effective.

This study attempted to evaluate the effects of three different conditioning procedures on the rate of vocalizations exhibited by children with a minimal or emerging echoic repertoire. We evaluated the effects of these conditioning procedures by comparing each within and across participants using a multielement and reversal design, along with a multiple baseline design. This study extends the literature on using conditioning procedures to increase vocalizations in several ways. First, it is the first comparison of all three conditioning procedures within the same participant. With respect to vocalizations, researchers have compared the effects of SSP and ODT within the same participant (i.e., Lepper et al., 2013) and the effects of RSP and SSP within the same participant (Lepper and Petursdottir 2017). However, no one has compared the effects of SSP, RSP, and ODT within the same participant. Therefore, it is unclear under what situations (e.g., participant characteristics) each conditioning procedure is most appropriate. By comparing all three procedures within the same participant, we had hoped to gather information concerning when one conditioning procedure may be more successful than another.

Second, the current study extends the literature by lending further support for the use of SSP, RSP, and ODT to increase vocalizations. With respect to SSP, current research findings are mixed, and researchers suggest the mixed results may be due to procedural variations across studies (Shillingsburg et al., 2015). In the current study, we conducted SSP according to the procedures suggested to be most effective (Esch, et al., 2009; Rader et al., 2014). However, SSP was only effective for two of our participants. Therefore, there may be other procedural variations that need to be considered. With respect to RSP and ODT, current research findings are limited, with only one publication demonstrating the effects of each conditioning procedure. In the current study, we conducted procedures similar to those of Lepper and Petursdottir (2017) for RSP and similar to those of Lepper et al. (2013) for ODT, but also incorporated the procedures suggested to be most effective for SSP (Esch, et al., 2009; Rader et al., 2014). Again, however, RSP was only effective for three participants and ODT was only effective for two participants, suggesting further investigation into the variables that may affect these procedures is warranted.

Overall, the results of the current study may have important implications for clinicians who work with individuals with delayed communication skills, specifically those who exhibit minimal vocalizations. Specifically, the current methods for teaching children to communicate (e.g., echoic training, mand training, reinforcement of vocalizations) are not always effective (Ward et al., 2007). A better understanding of how and when to use SSP, RSP, and ODT may give clinicians more options for teaching communication. A better understanding of the most effective conditioning procedures may allow clinicians to more effectively and efficiently teach communication.

References

Benasich, A. A. & Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioral Brain Research, 136,* 31-49. doi: 10.1016/S0166-4328(02)00098-0

Bijou, S. W & Baer, D. M. (1965). *Child development volume II: Universal state of infancy.* New York, NY: Appleton-Century-Crofts.

Carroll, R. A. & Klatt, K. P. (2008). Using stimulus-stimulus pairing and direct reinforcement to teach vocal verbal behavior to young children with autism. *The Analysis of Verbal Behavior, 24,* 135-146.

Dinsmoor, J. A. (1995a). Stimulus control: Part I. *The Behavior Analyst, 18,* 51-68.

Dinsmoor, J. A. (1995b). Stimulus control: Part II. *The Behavior Analyst, 18,* 253-269.

Dorow, L. G. (1980). Generalization effects of newly conditioned reinforcers. *Education and Training of the Mentally Retarded, 15,* 8-14.

Dozier, C. L., Iwata, B. A., Thomason-Sassi, J., Worsdell, A. S., & Wilson, D. M. (2012). A comparison of two pairing procedures to establish praise as a reinforcer. *Journal of Applied Behavior Analysis, 45,* 721-735. doi: 10.1901/jaba.2012.45-721

Esch, B. E. (2008). EESA: Early echoic skills assessment. In M. L. Sundberg VB-MAPP: Verbal behavior milestones assessment and placement program (pp. 62-63). Concord, CA: AVB Press.

Esch, B. E., Carr, J. E., & Grow, L. L. (2009). Evaluation of an enhanced stimulus-stimulus pairing procedure to increase early vocalizations of children with autism. *Journal of Applied Behavior Analysis, 42,* 225-241. doi: 10.1901/jaba.2009.42-225

Esch, B. E., Carr, J. E., & Michael, J. (2005) Evaluating stimulus-stimulus pairing and direct reinforcement in the establishment of an echoic repertoire of children diagnosed with autism. *The Analysis of Verbal Behavior, 21,* 43-58

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences, 27,* 491-541. doi: 10.1017/S0140525X04000111

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491-498. doi: 10.1901/jaba.1992.25-491

Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Effects of conditioning voices as reinforcers for listener responses on rate of learning, awareness, and preferences for listening to stories in preschoolers with autism. *The Analysis of Verbal Behavior, 27,* 103-124.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30,* 509-516. doi: 10.1177/0165025406071914

Hains, A. H. & Baer, D. M. (1989). Interaction effects in multielement designs: inevitable, desirable, and ignorable. *Journal of Applied Behavior Analysis, 22,* 57-69. doi: 10.1901/jaba.1989.22-57

Hall, G. & Sundberg, M. L. (1987). Teaching mands by manipulating conditioned establishing operations. *The Analysis of Verbal Behavior, 5*, 41-53.

Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children.* Baltimore, MD: Brookes

Hoff, E. (2005). *Language development (3rd ed.).* Belmont, CA: Wadsworth.

Inglis, I. R., Forkman, B., & Lazarus, J. (1997). Free food or earned food? A review and fuzzy model of contrafreeloading. *Animal Behavior, 53,* 1171-1191. doi: 10.1006/anbe.1996.0320

Kjelgaard, M. M. & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16,* 287-308. doi: 10.1080/01690960042000058

Lepper, T. L. & Petursdottir, A. I. (2017). Effects of response-contingent stimulus pairing on vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis, 50,* 756-774. doi: 10.1002/jaba.415

Lepper, T. L., Petursdottir, A. I., & Esch, B. E. (2013). Effects of operant discrimination training on the vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis, 46,* 656-661. doi: 10.1002/jaba.55

Lovaas, O. I. (2003). *Teaching Individuals with Developmental Delays.* Austin, Texas: Pro-Ed.

Lovaas, O. I., Freitag, G., Kinder, M. I., Rubenstein, B. D., Schaeffer, B., & Simmons, J. Q. (1966). Establishment of social reinforcers in two schizophrenic children on the basis of food. *Journal of Experimental Child Psychology, 4,* 109-125. doi: 10.1016/0022-0965(66)90011-7

Lucyznski, K. C. & Hanley, G. P. (2009). Do children prefer contingencies? An evaluation of the efficacy of and preference for contingent versus noncontingent social attention. *Journal of Applied Behavior Analysis, 42,* 511-525. doi: 10.1901/jaba. 2010.43-397

Miguel, C. F., Carr, J. E., & Micahel, J. (2002). The effects of a stimulus-stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior, 18,* 3-13.

Miliotis, A., Sidener, T. M., Reeve, K. F., Carbone, V., Sidener, D. W., Rader, L., & Delmolino, L. (2012). An evaluation of the number of presentations of target sounds during stimulus-stimulus pairing trials. *Journal of Applied Behavior Analysis, 45,* 809-813. doi: 10.9101/jaba.2012.45-809

Normand, M. P. & Knoll, M. L. (2006). The effects of a stimulus-stimulus pairing procedure on the unprompted vocalizations of a young child diagnosed with autism. *The Analysis of Verbal Behavior, 22,* 81-85.

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18,* 249-255. doi: 10.1901/jaba.1985.18-249

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders, 44,* 2413-2428. doi: 10.1007/s10803-014-2047-4

Peña-Brooks, A. & Hegde, M. N. (2007). Assessment and Treatment of Articulation and Phonological Disorders in Children. Austin, TX: Pro-Ed.

Petursdottir, A. I., Carp, C. L., Matthies, D. W., & Esch, B. E. (2011). Analyzing stimulus-stimulus pairing effects on preferences for speech sounds. *The Analysis of Verbal Behavior, 27,* 45-60.

Petursdottir, A. I. & Mellor, J. R. (2017). Reinforcement contingencies in language acquisition: Implications for language intervention. *Policy Insights from the Behavioral and Brain Sciences, 4,* 25-32. doi: 10.1177/2372732216686083

Petursdottir, A. I. & Lepper, T. L. (2015). Inducing novel vocalizations by conditioning speech sounds as reinforcers. *Behavior Analysis in Practice, 8,* 223-232. doi: 10.1007/s40617-015-0088-6

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders, 44,* 2413-2428. doi: 10.1007/s10803-014-2047-4

Rader, L., Sidener, T. M. Reeve, K. F., Sidener, D. W., Delmolino, L., Miliotis, A., & Carbone, V. (2014). Stimulus-stimulus pairing and vocalizations: A systematic replication. *The Analysis of Verbal Behavior, 30,* 69-74. doi: 10.1007/s40616-014-0012-0

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151-160. doi: 10.1037/0003-066X.43.3.151

Ross D. E. & Greer, R. D. (2003). Generalized imitation and the mand: inducing first instances of speech in young children with autism. *Research in Developmental Disabilities, 24,* 58-74. doi: 10.1016/S0891-4222(02)00167-1

Schlinger, H. D. (1995). *A behavior analytic view of child development.* New York, NY: Plenum Press.

Shillingsburg, M. A., Hollander, D. L., Yosick, R. N., Bowen, C., & Muskat, L. R. (2015). Stimulus-stimulus pairing to increase vocalizations in children with language delays: A review. *The Analysis of Verbal Behavior, 31,* 215-235. doi: 10.1007/s40616-015-0042-2

Sidowski, J. B., Kass, N., & Wilson, H. (1965). Cue and secondary reinforcement effects with children. *Journal of Experimental Psychology, 69,* 340-342. doi: 10.1037/h0021785

Skinner, B. F. (1938). *The Behavior of Organisms.* Acton, MA: Copley Publishing Group.

Skinner, B. F. (1957). *Verbal Behavior.* Cambridge, MA: BF Skinner Foundation.

Skinner, B. F. (1969). *Contingencies of Reinforcement.* Cambridge, MA: BF Skinner Foundation.

Skinner, B. F. (1982). Contrived reinforcement. *The Behavior Analyst, 5,* 3-8.

Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcement and automatic punishment in infant vocal behavior. *The Analysis of Verbal Behavior, 13,* 39-48.

Stock, R. A., Schulze, K. A., & Mirenda, P. (2008). A comparison of stimulus-stimulus pairing, standard echoic training, and control procedures on the vocal behavior of children with autism. *The Analysis of Verbal Behavior, 24,* 123-133.

Sundberg M. L. VB-MAPP (2008): *Verbal behavior milestones assessment and placement program*. Concord, CA: AVB Press

Sundberg, M. L. & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25,* 698-724. doi: 10.1177/0145445501255003

Sundberg, M. L., Michael, J., Partington, J. W., & Sundberg, C.A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior, 13,* 21-37.

Taylor-Santa, C., Sidener, T. M., Carr, J. E., & Reeve, K. F. (2014). A discrimination training procedure to establish conditioned reinforcers for children with autism. *Behavioral Interventions, 29,* 157-176. doi: 10.1002/bin.1384

Toth, K., Munson, J., Meltzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders, 36,* 993-1005. doi: 10.1007/s10503-006-0137-7

Vaughan, M. E. & Michael, J. L. (1982). Automatic reinforcement: An important but ignored concept. *Behaviorism, 10,* 217-227.

Ward. S. J., Osnes, P. J., & Partington, J. W. (2007). The effects of a delay of noncontingent reinforcement during a pairing procedure in the development of stimulus control of automatically-reinforced vocalizations. *The Analysis of Verbal Behavior, 23,* 103-111.

Yoon, S. & Bennett, G. M. (2000). Effects of a stimulus-stimulus pairing procedure on conditioning vocal sounds as reinforcers. *The Analysis of Verbal Behavior, 17,* 75-88.

Yoon, S. & Feliciano, G. M. (2007). Stimulus-stimulus pairing and subsequent mand acquisition of children with various levels of verbal repertoires. *The Analysis of Verbal Behavior, 23,* 3-16.

Figure Captions

Tables

Table 1. Demographic characteristics of participant sample

Characteristics		Number of subjects	Percent of sample
Gender	Male	3	43%
	Female	4	57%
Age	12-24 months	6	86%
	25-36 months	1	14%
Classroom Setting	Inclusive Education	4	57%
	Early Intensive Behavioral Intervention (EIBI)	3	43%
Diagnosis	None	3	43%
	Autism	2	29%
	Down Syndrome	1	14%
	Hypoxic-ischemic Encephalopathy; Global Developmental Delay; Cerebral Palsy	1	14%
Early Echoic Skills	0-5	6	86%
Assessment Score	15-20	1	14%

Table 2. Pre- and Post-conditioning EESA Scores

Participants	Pre-conditioning EESA Score	Post-conditioning EESA Score
Benny (Sound Set 2)	N/A	3.5
Cynthia	17.5	18.5
Benny (Sound Set 1)	0	N/A
Allan	0	30.5
Octavia	0	18.5
Maisy	0	0
Mason	0	0
Antonio	0	0

Table 3. Pre-existing Skills Repertoire

Participants	Assessment Type	Total Score	Listener Skills	Visual Perceptual & Match-to-Sample	Imitation	Vocal
Benny	VB-MAPP	17	2	3.5	2.5	3
Cynthia	N/A	N/A	N/A	N/A	N/A	N/A
Allan	N/A	N/A	N/A	N/A	N/A	N/A
Mason	VB-MAPP	19	1	3.5	0.5	1
Antonio	VB-MAPP	18	1	4	0	1
Maisy	Carolina	24	0	0	0	0
Octavia	N/A	N/A	N/A	N/A	N/A	N/A

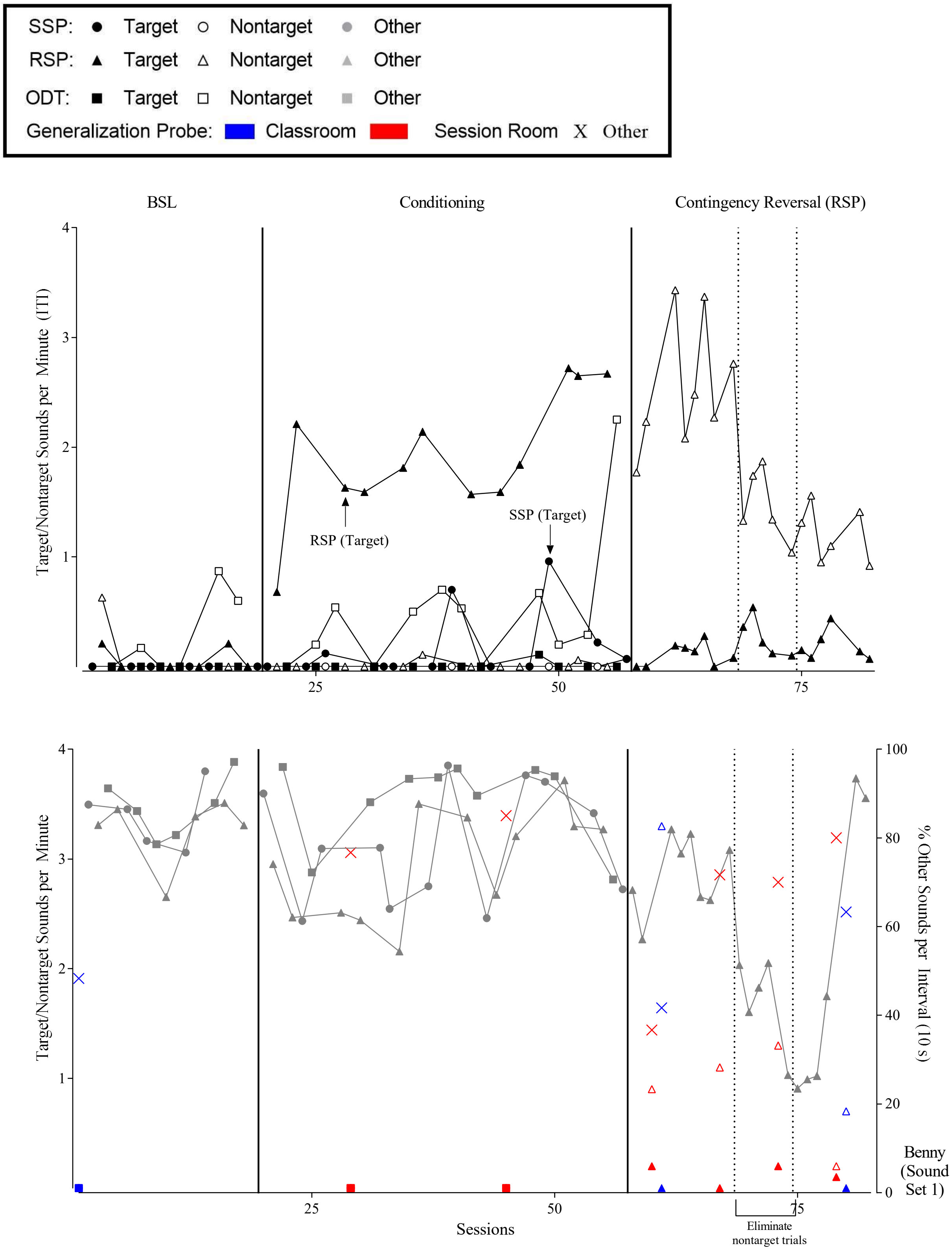
Table 4. Summary of Results (Study 1)

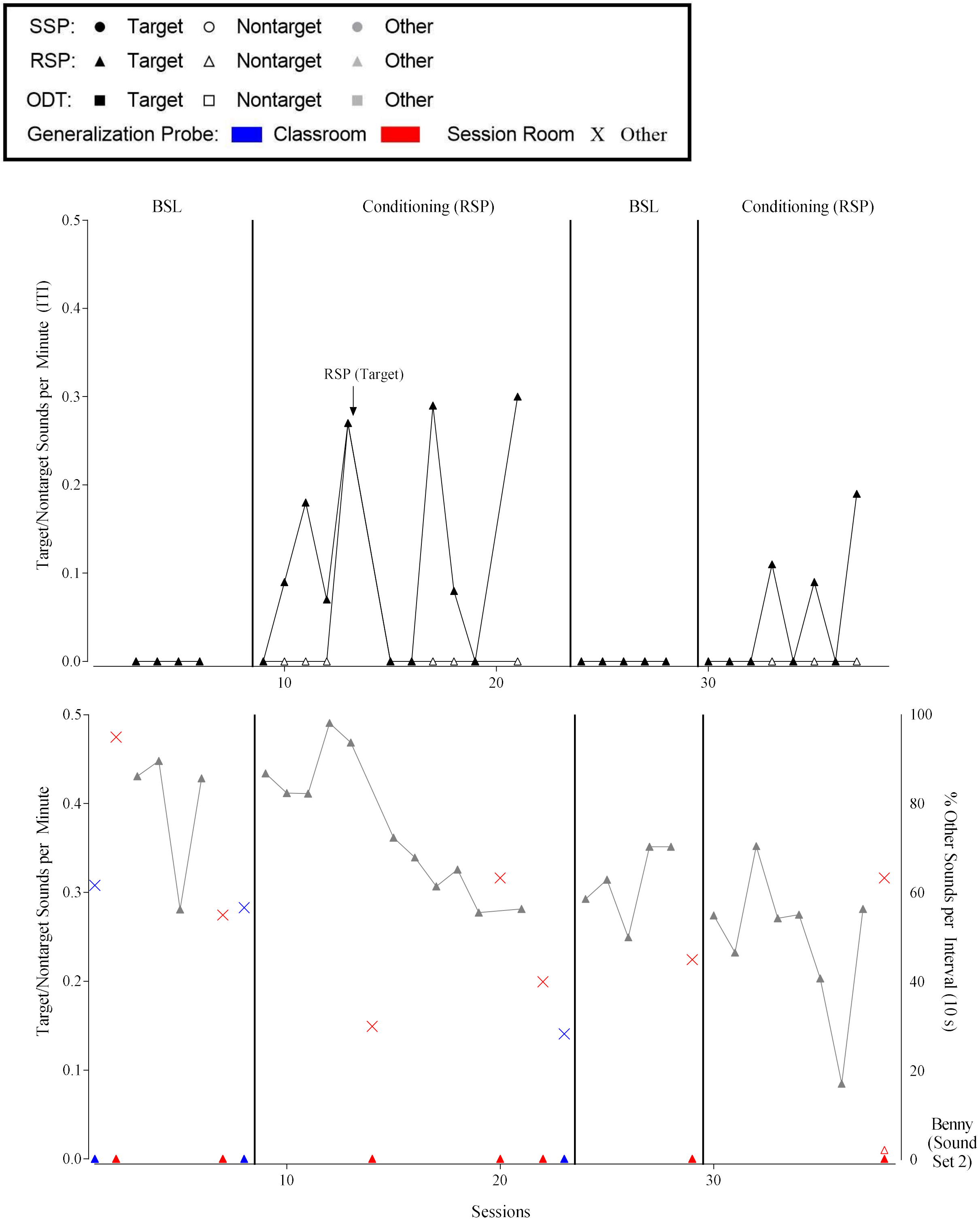
Participants	Increase *During* Conditioning (ITI)				Increase *Following* Conditioning	
	SSP	RSP	ODT	None	Generalization	EESA Score
Benny (Sound Set 2)	N/A	Y	N/A	N	N	Y
Cynthia	Y	Y	Y	N	Y	Y
Benny (Sound Set 1)	Y	Y	N	N	Y	N/A
Allan	N	N	Y	N	Y	Y
Octavia	N	N	N	Y*	Y	Y
Maisy	N	N	N	Y	N	N
Mason	N	N	N	Y	N	N
Antonio	N	N	N	Y	N	N
Totals:	2/7 (29%)	3/8 (38%)	2/7 (29%)	4/8 (50%)	4/8 (50%)	4/7 (57%)

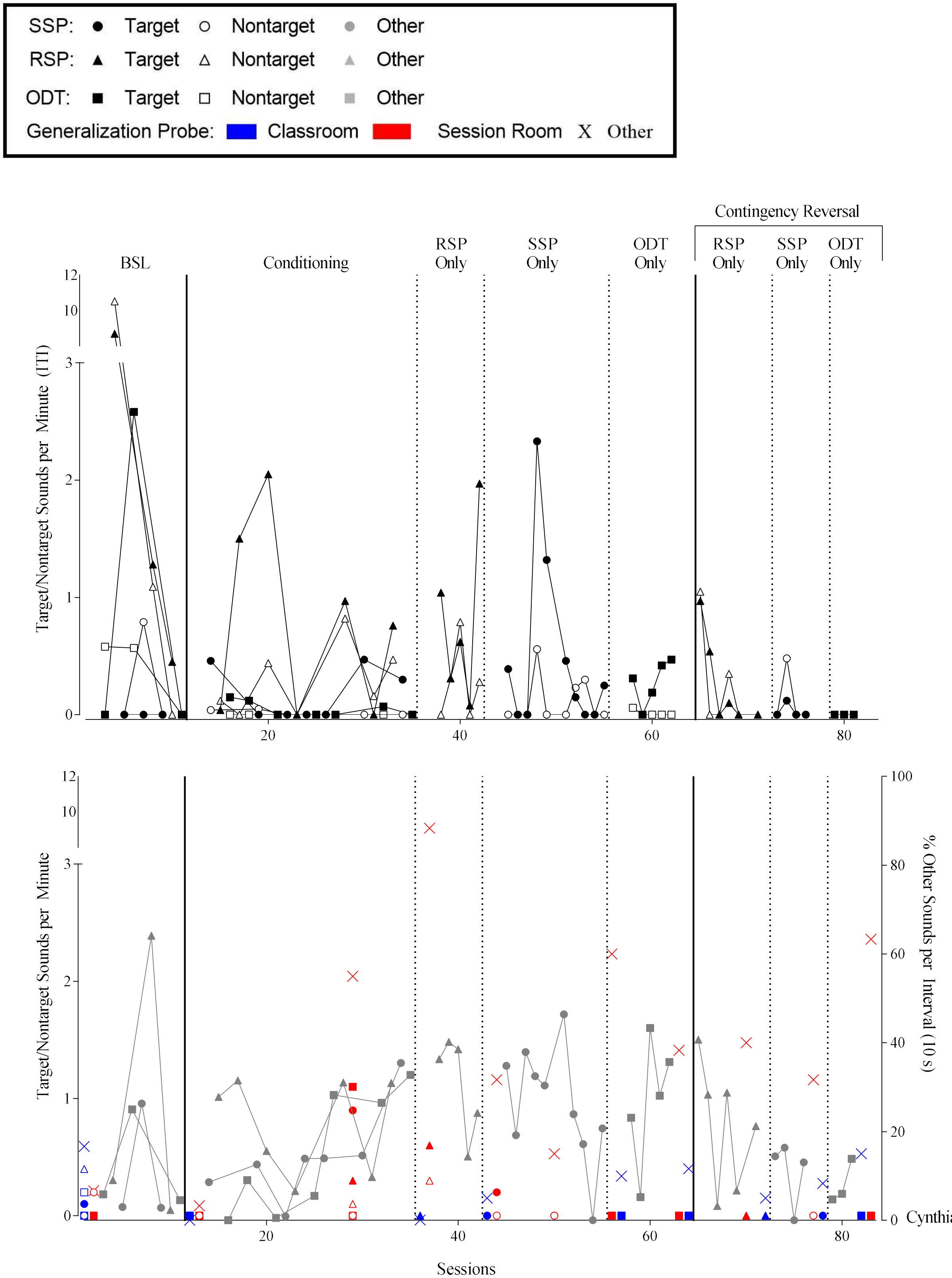
*Increased during trial – suggesting echoic repertoire emerged

Table 5. Summary of Results (Study 2)

	Barriers		
Participants	Articulation	Attending	Reinforcer Efficacy
Maisy	N	N	Y
Mason	Y	N	N/A
Antonio	N	Y	N/A
Totals:	1/2 (50%)	1/2 (50%)	1/1 (100%)


SSP: Target Nontarget Other
RSP: Target Nontarget Other
ODT: Target Nontarget Other
Generalization Probe: Classroom Session Room X Other
BSL
Conditioning
Contingency Reversal (RSP)
Target/Nontarget Sounds per Minute (ITI)
RSP (Target)
SSP (Target)
25
50
75
Target/Nontarget Sounds per Minute
% Other Sounds per Interval (10 s)
Benny (Sound Set 1)
Sessions
Eliminate nontarget trials


SSP: Target Nontarget Other
RSP: Target Nontarget Other
ODT: Target Nontarget Other
Generalization Probe: Classroom Session Room X Other
BSL
Conditioning (RSP)
BSL
Conditioning (RSP)
RSP (Target)
Target/Nontarget Sounds per Minute (ITI)
Target/Nontarget Sounds per Minute
% Other Sounds per Interval (10 s)
Benny (Sound Set 2)
Sessions


SSP: Target Nontarget Other
RSP: Target Nontarget Other
ODT: Target Nontarget Other
Generalization Probe: Classroom Session Room X Other
Contingency Reversal
BSL
Conditioning
RSP Only
SSP Only
ODT Only
RSP Only
SSP Only
ODT Only
Target/Nontarget Sounds per Minute (ITI)
Target/Nontarget Sounds per Minute
% Other Sounds per Interval (10 s)
Sessions
Cynthia

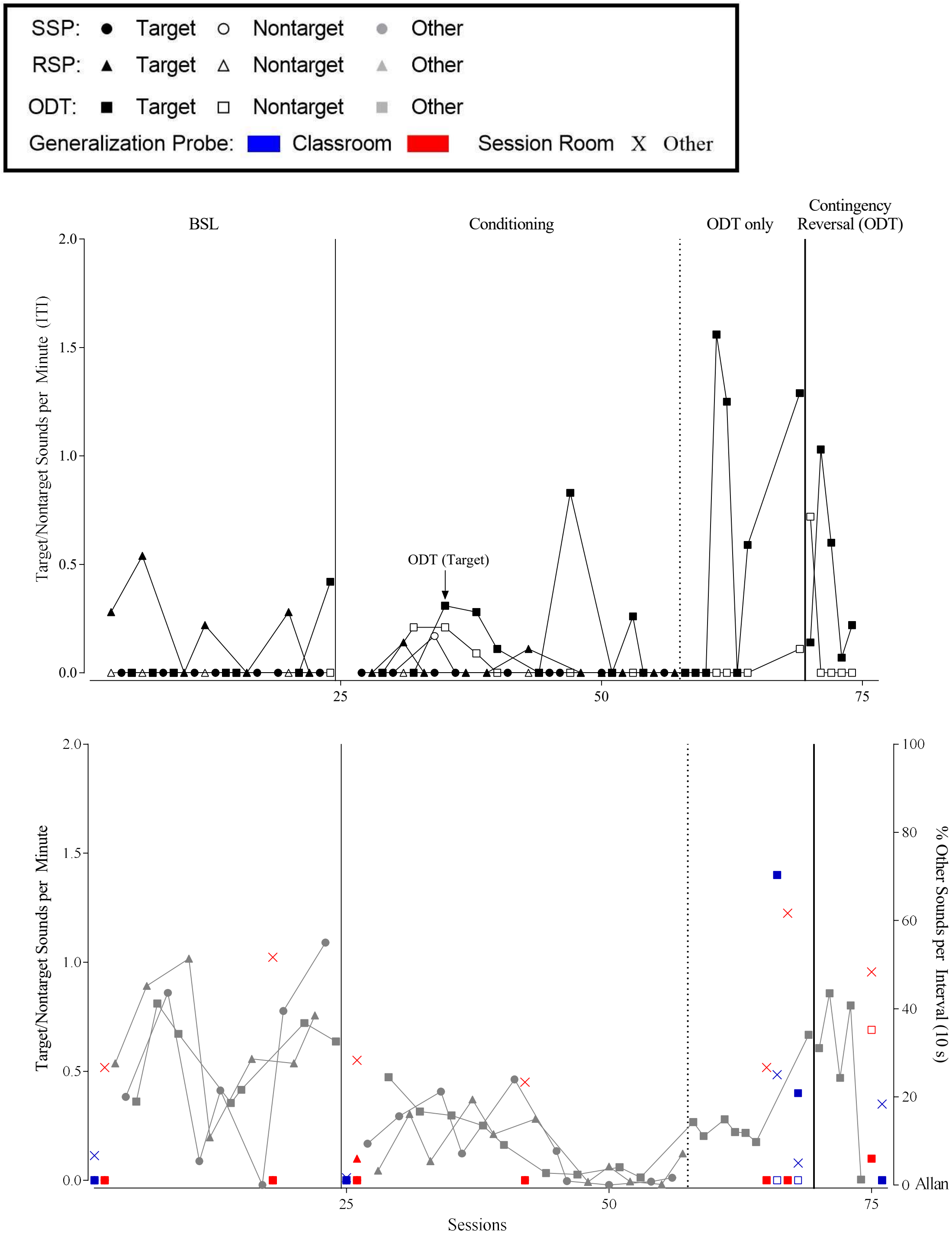
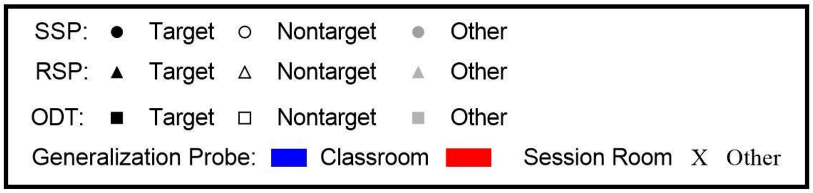
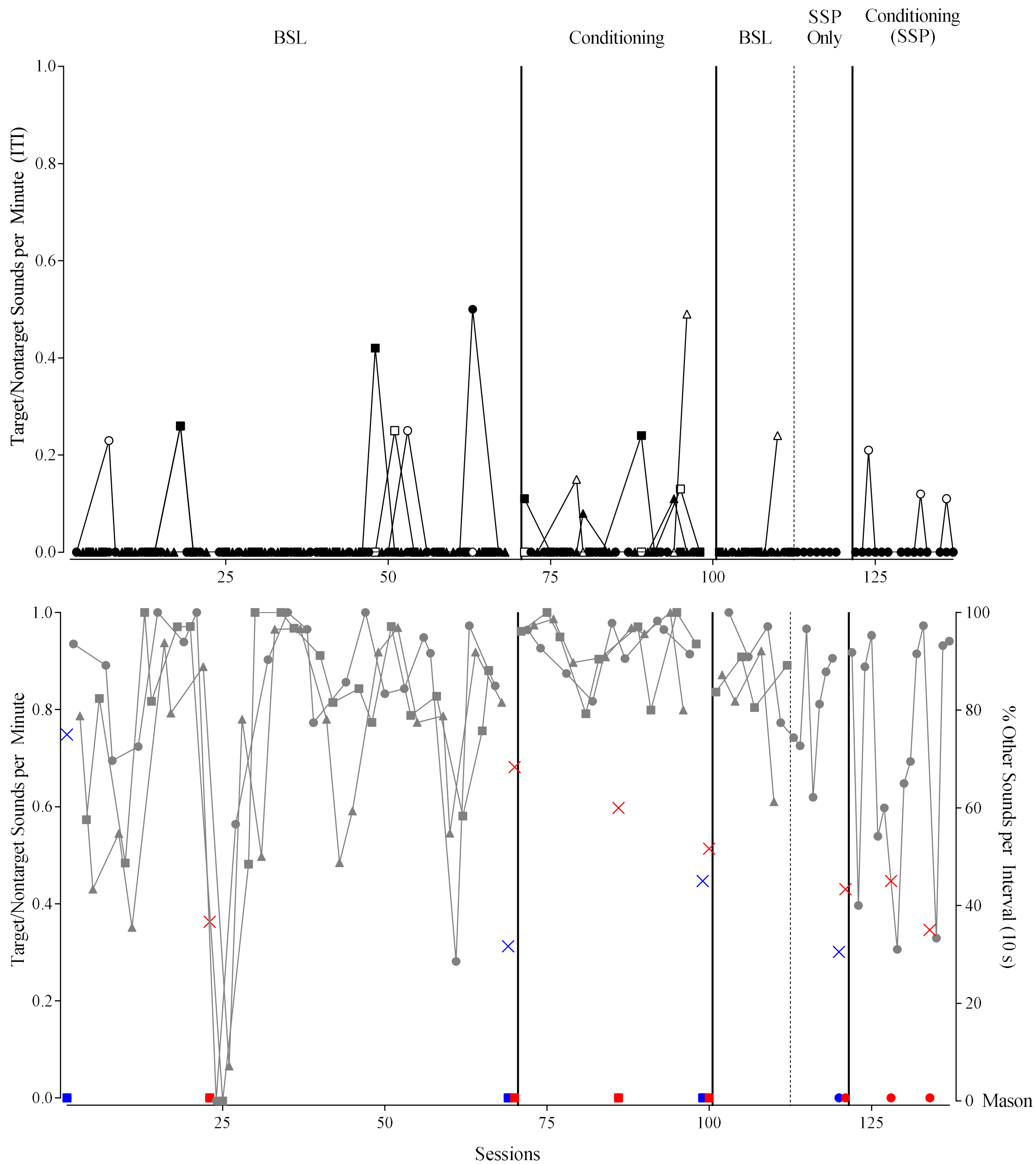
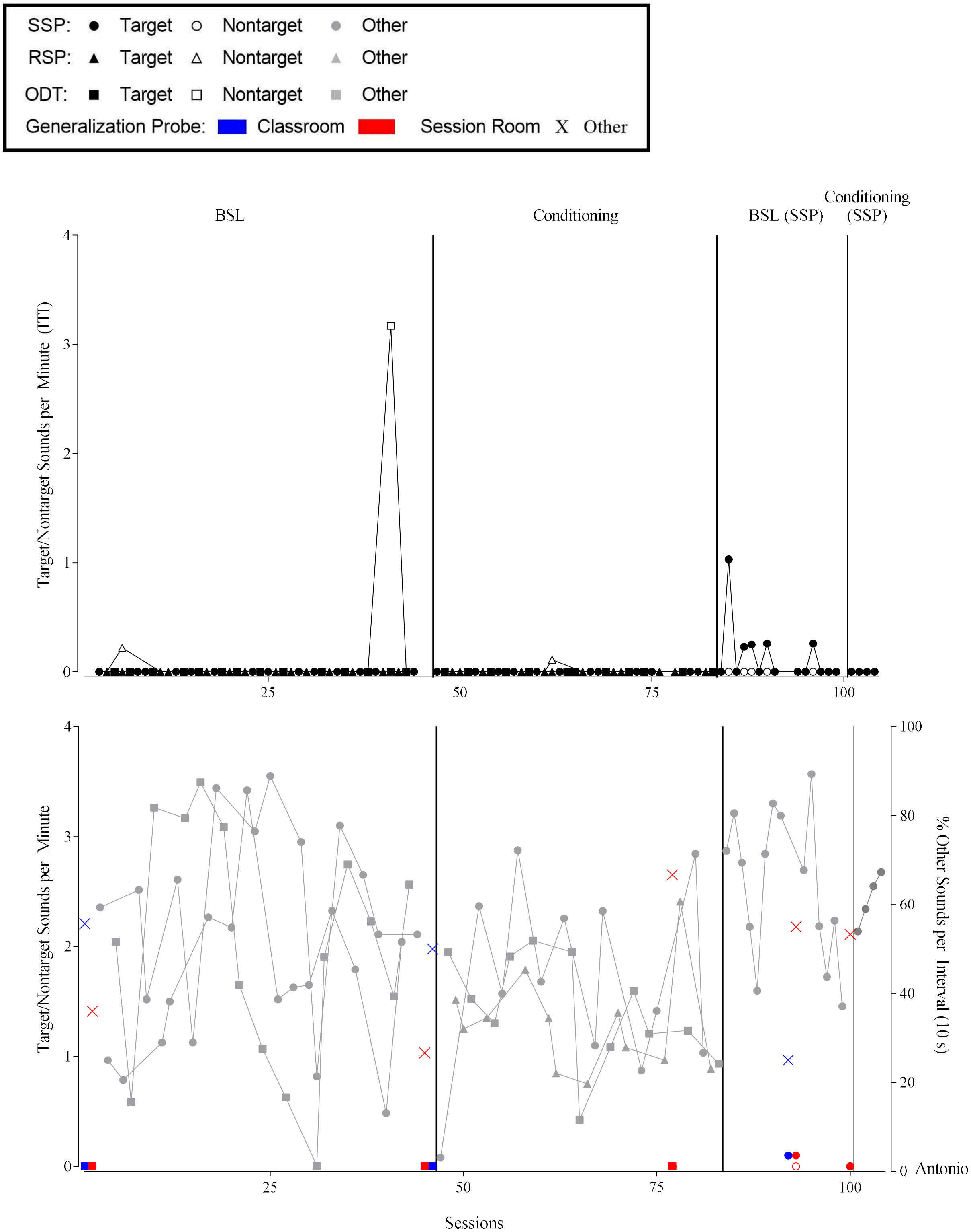


Figure 1: Rate of *novel* target and nontarget vocalizations during the ITI and percentage of intervals with other vocalizations for each participant across stimulus-stimulus pairing (SSP), response-stimulus pairing (RSP), and operant discrimination training (ODT). The break in the y-axis and data path for Cynthia was inserted to accommodate an outlier during RSP, session 2.


SSP: Target Nontarget Other
RSP: Target Nontarget Other
ODT: Target Nontarget Other
Generalization Probe: Classroom Session Room X Other


BSL
Conditioning
BSL
SSP Only
Conditioning (SSP)
Target/Nontarget Sounds per Minute (ITI)
Target/Nontarget Sounds per Minute
% Other Sounds per Interval (10 s)
Sessions
Mason


SSP: Target Nontarget Other
RSP: Target Nontarget Other
ODT: Target Nontarget Other
Generalization Probe: Classroom Session Room X Other
BSL
Conditioning
BSL (SSP)
Conditioning (SSP)
Target/Nontarget Sounds per Minute (ITI)
Target/Nontarget Sounds per Minute
% Other Sounds per Interval (10 s)
Antonio
Sessions

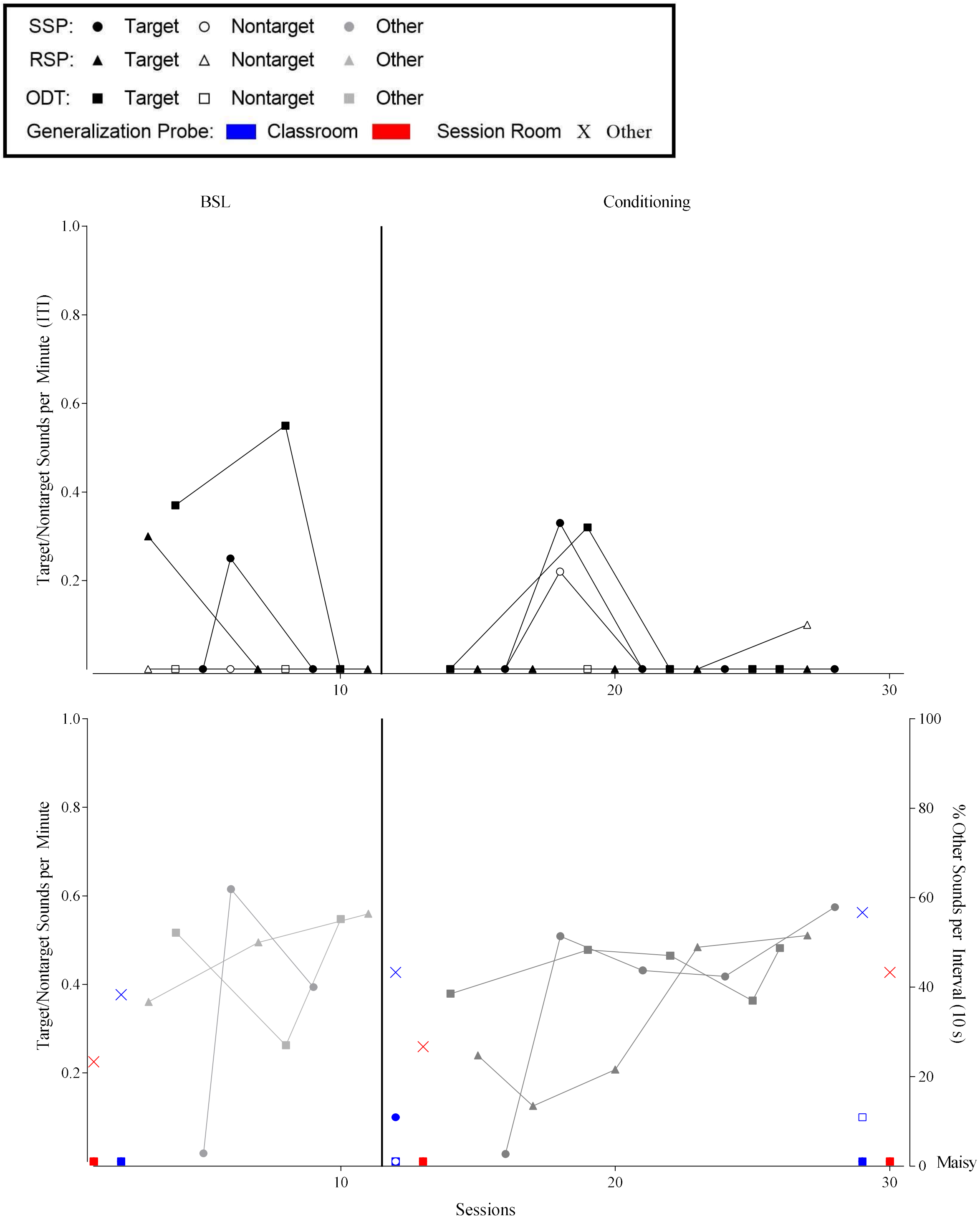
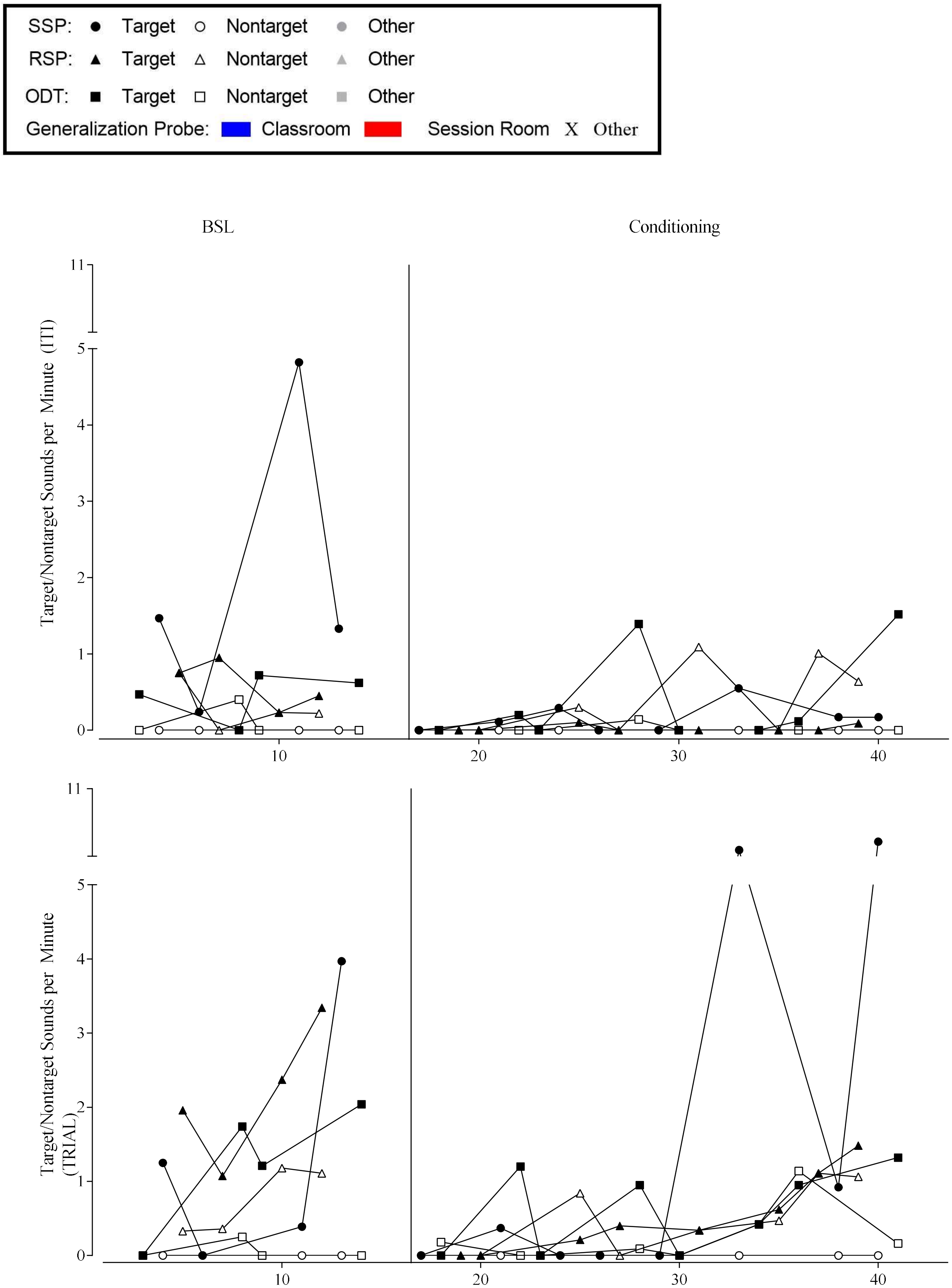


Figure 2: Rate of *novel* target and nontarget vocalizations during the ITI and percentage of intervals with other vocalizations for each participant across stimulus-stimulus pairing (SSP), response-stimulus pairing (RSP), and operant discrimination training (ODT).


SSP: Target Nontarget Other
RSP: Target Nontarget Other
ODT: Target Nontarget Other
Generalization Probe: Classroom Session Room X Other
BSL
Conditioning
Target/Nontarget Sounds per Minute (ITI)
Target/Nontarget Sounds per Minute (TRIAL)
11
5
4
3
2
1
0
10
20
30
40

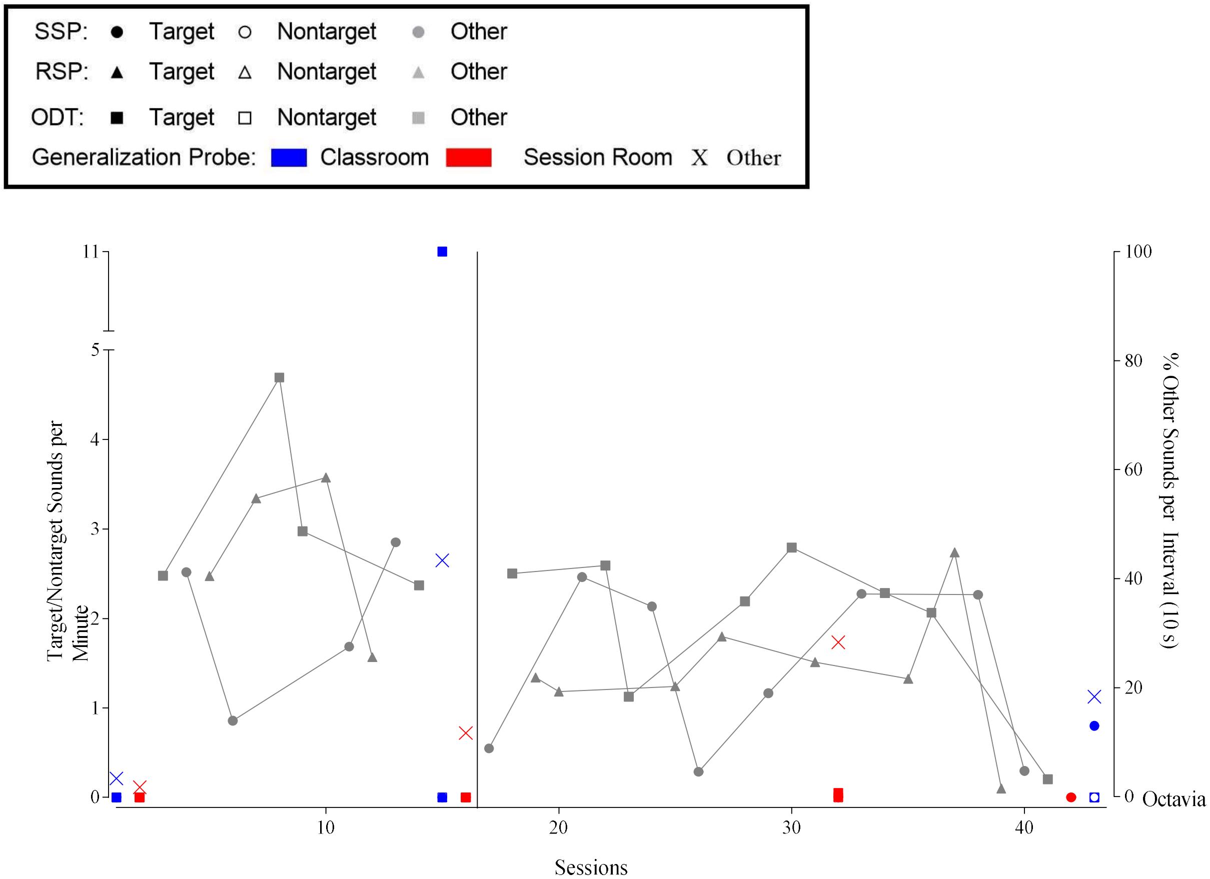


Figure 3: Rate of *novel* target and nontarget vocalizations during the ITI (top graph) and during the trial (middle graph) and percentage of intervals with other vocalizations for Octavia across stimulus-stimulus pairing (SSP), response-stimulus pairing (RSP), and operant discrimination training (ODT). The break in the y-axis and data path for Octavia was inserted to accommodate an outlier during the classroom generalization probe, session 15 (top graph) and during SSP, session 33 and 39.

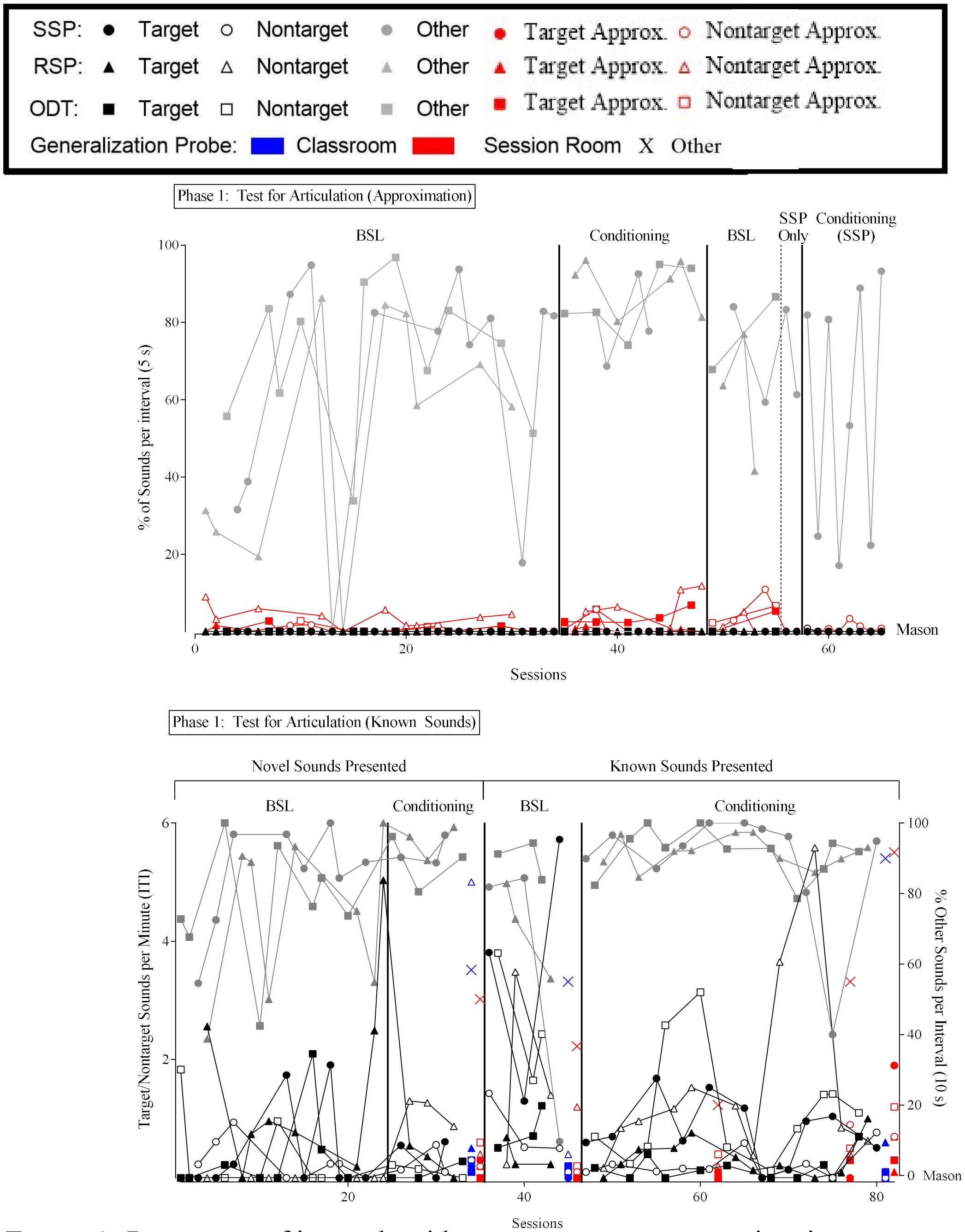


Figure 4: Percentage of intervals with target, nontarget, approximations to target and nontarget, and other vocalizations (top graph) and rate of *known* target and nontarget vocalizations during the ITI and percentage of intervals with other vocalizations (bottom graph) for Mason across stimulus-stimulus pairing (SSP), response-stimulus pairing (RSP), and operant discrimination training (ODT).

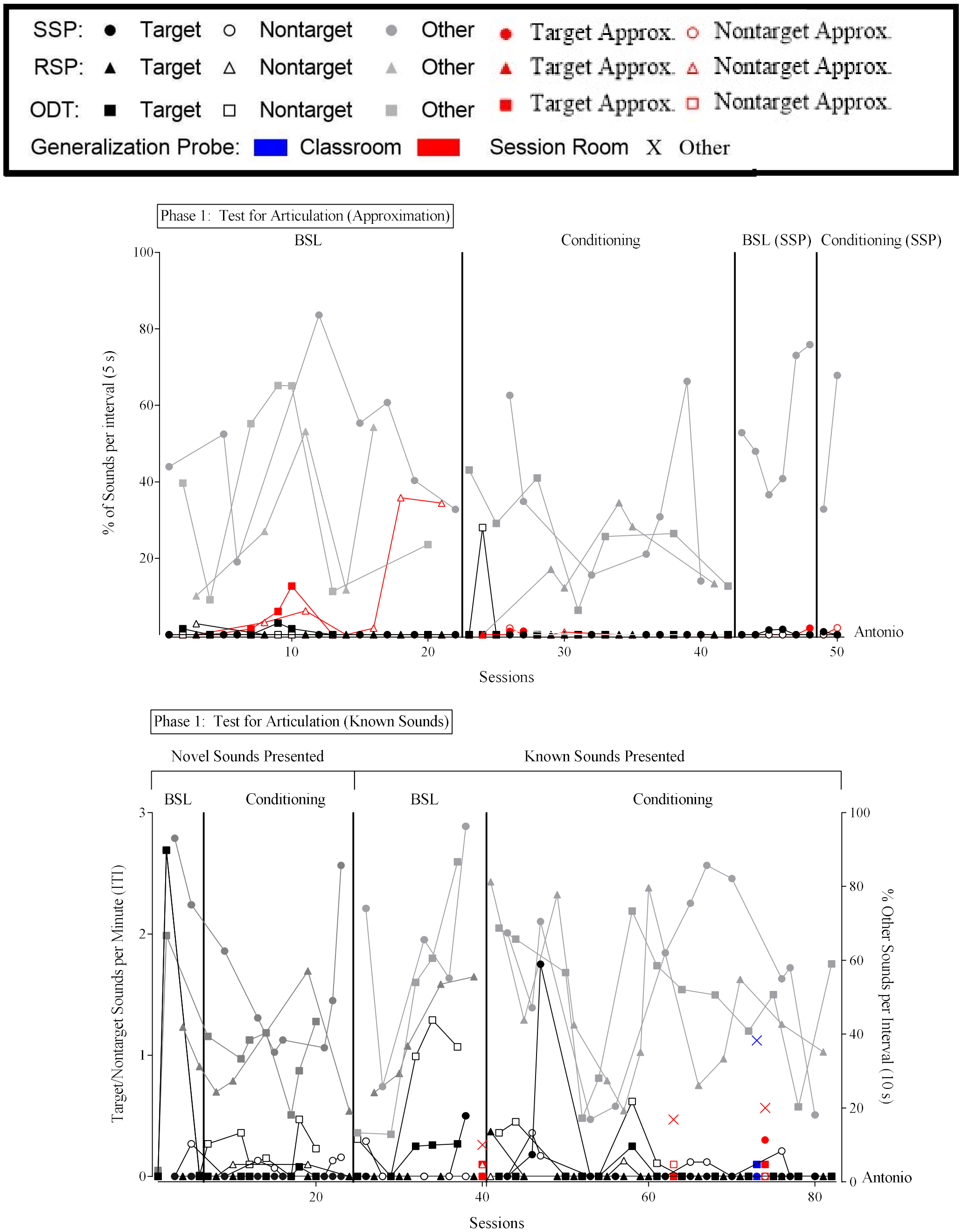


Figure 5: Percentage of intervals with target, nontarget, approximations to target and nontarget, and other vocalizations (top graph) and rate of *known* target and nontarget vocalizations during the ITI and percentage of intervals with other vocalizations (bottom graph) for Antonio across stimulus-stimulus pairing (SSP), response-stimulus pairing (RSP), and operant discrimination training (ODT).

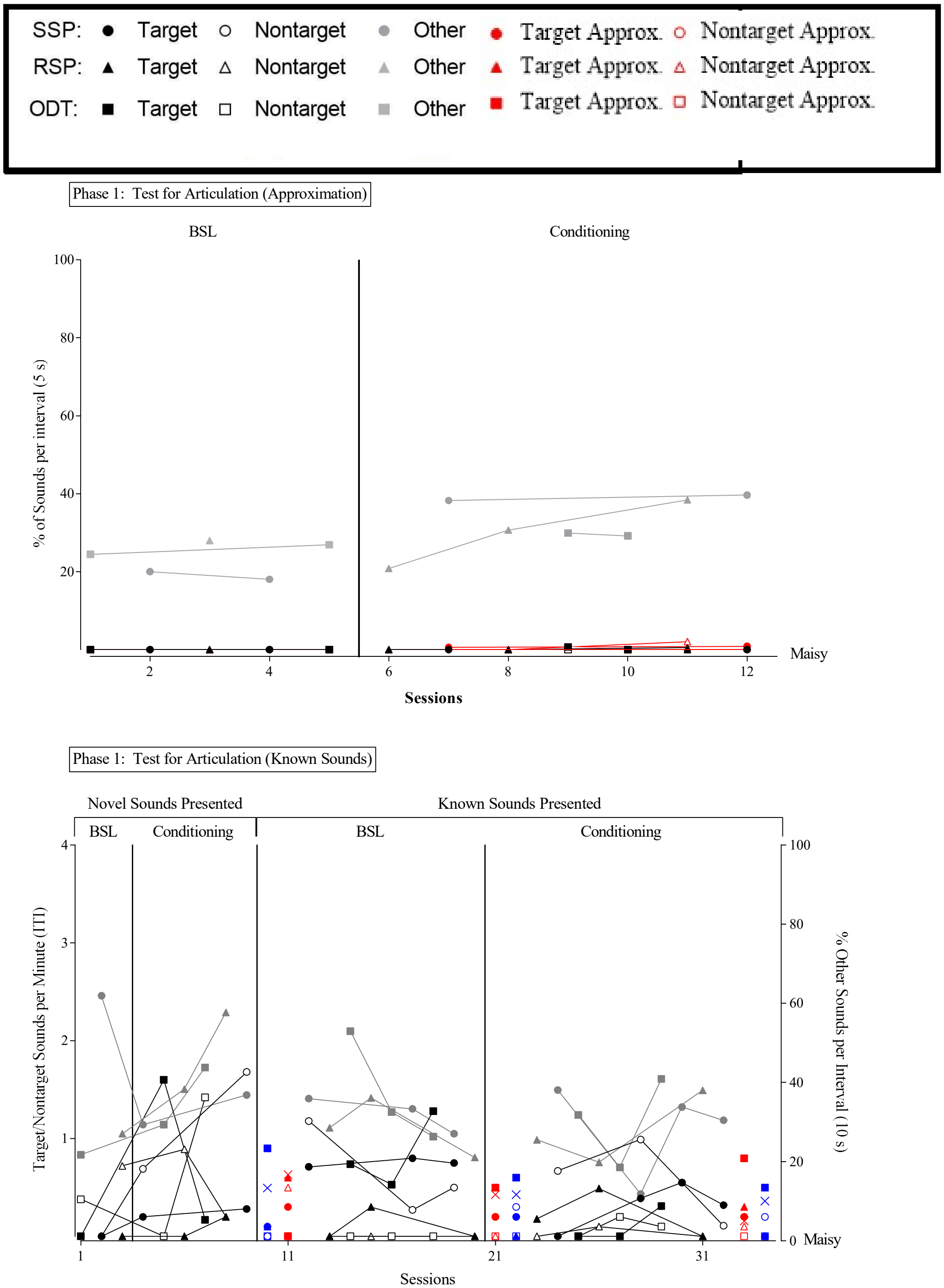


Figure 6: Percentage of intervals with target, nontarget, approximations to target and nontarget, and other vocalizations for Maisy across stimulus-stimulus pairing (SSP), response-stimulus pairing (RSP), and operant discrimination training (ODT).

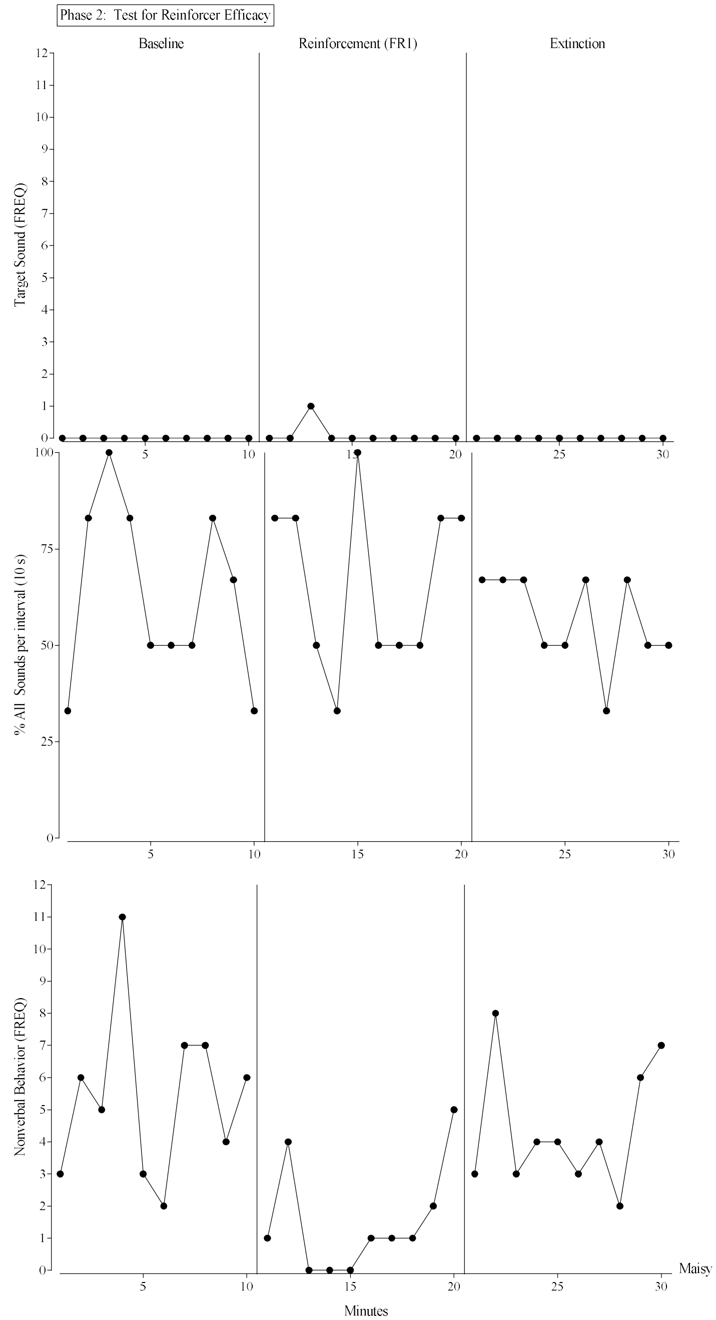


Figure 7: Frequency of target sounds (top graph), percentage of intervals with all sounds (middle graph), and frequency of a nonverbal behavior (bottom graph) for Maisy across three reinforcer assessments.

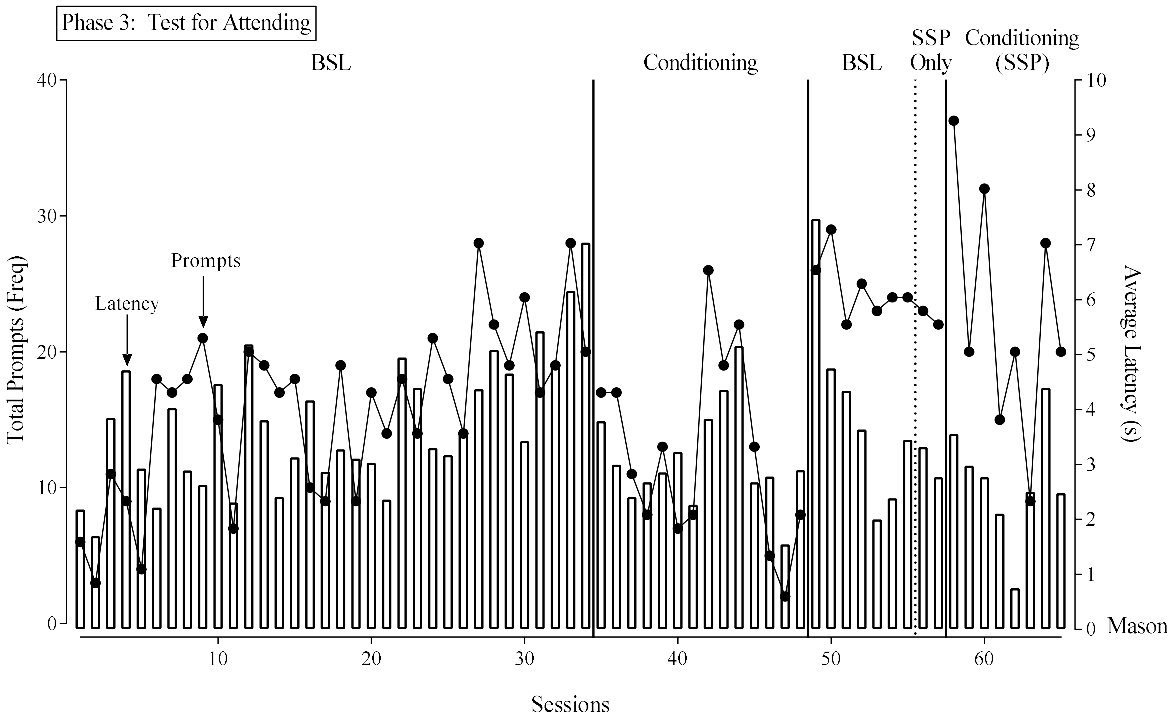


Figure 8: Frequency of total prompts and average latency for Mason across sessions.

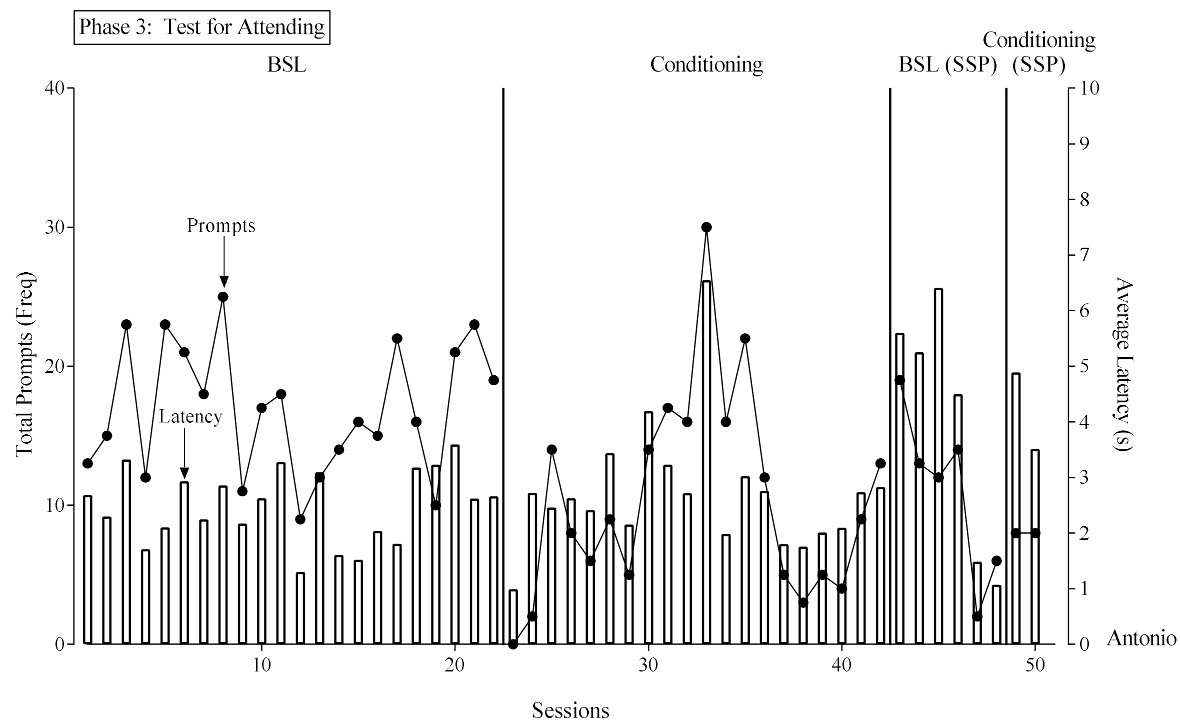
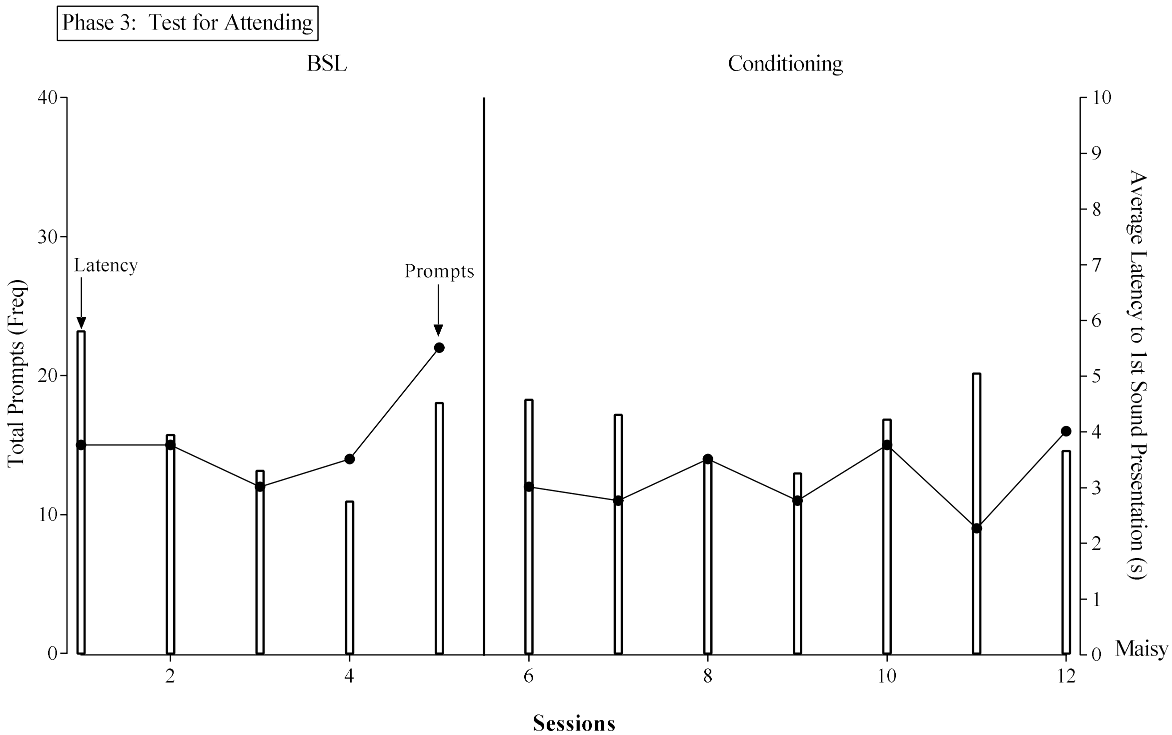


Figure 9: Frequency of total prompts and average latency for Antonio across sessions.



Phase 3: Test for Attending
BSL
Conditioning
Total Prompts (Freq)
Average Latency to 1st Sound Presentation (s)
Latency
Prompts
Maisy
Sessions

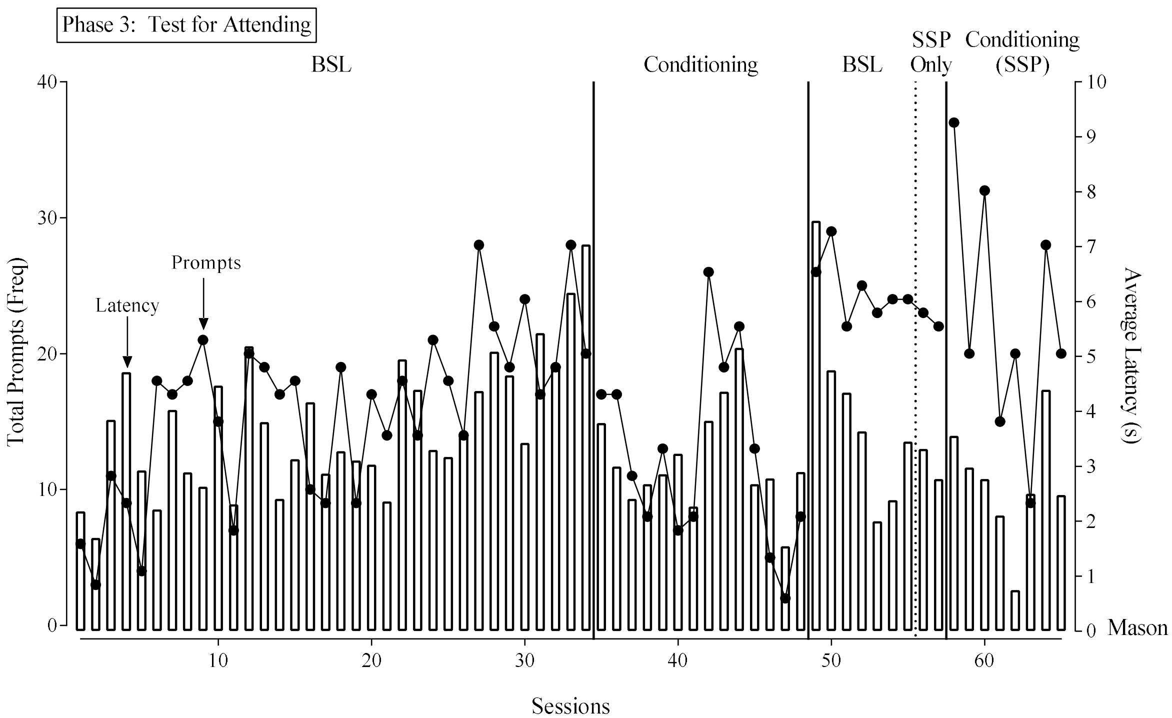


Figure 10: Frequency of total prompts and average latency for Maisy across sessions.

Appendix A

Early Echoic Skills Assessment (EESA)

Barbara E. Esch, Ph.D., BCBA, CCC-SLP

Scoring Groups 1-3: For each item, score the best response of up to 3 trials
X = correct sounds and correct number of syllables (1 point)
/ = recognizable response, but incorrect or missing consonants or extra syllables (½ point)
Blank = no response, incorrect vowels, or missing syllables (0 points)

TOTAL RAW SCORE: *(Groups 1-5)*

Assessment			
1st	2nd	3rd	4th

Group 1: Simple and reduplicated syllables

Targets: vowels, diphthongs, consonants p, b, m, n, h, w

Probe: t

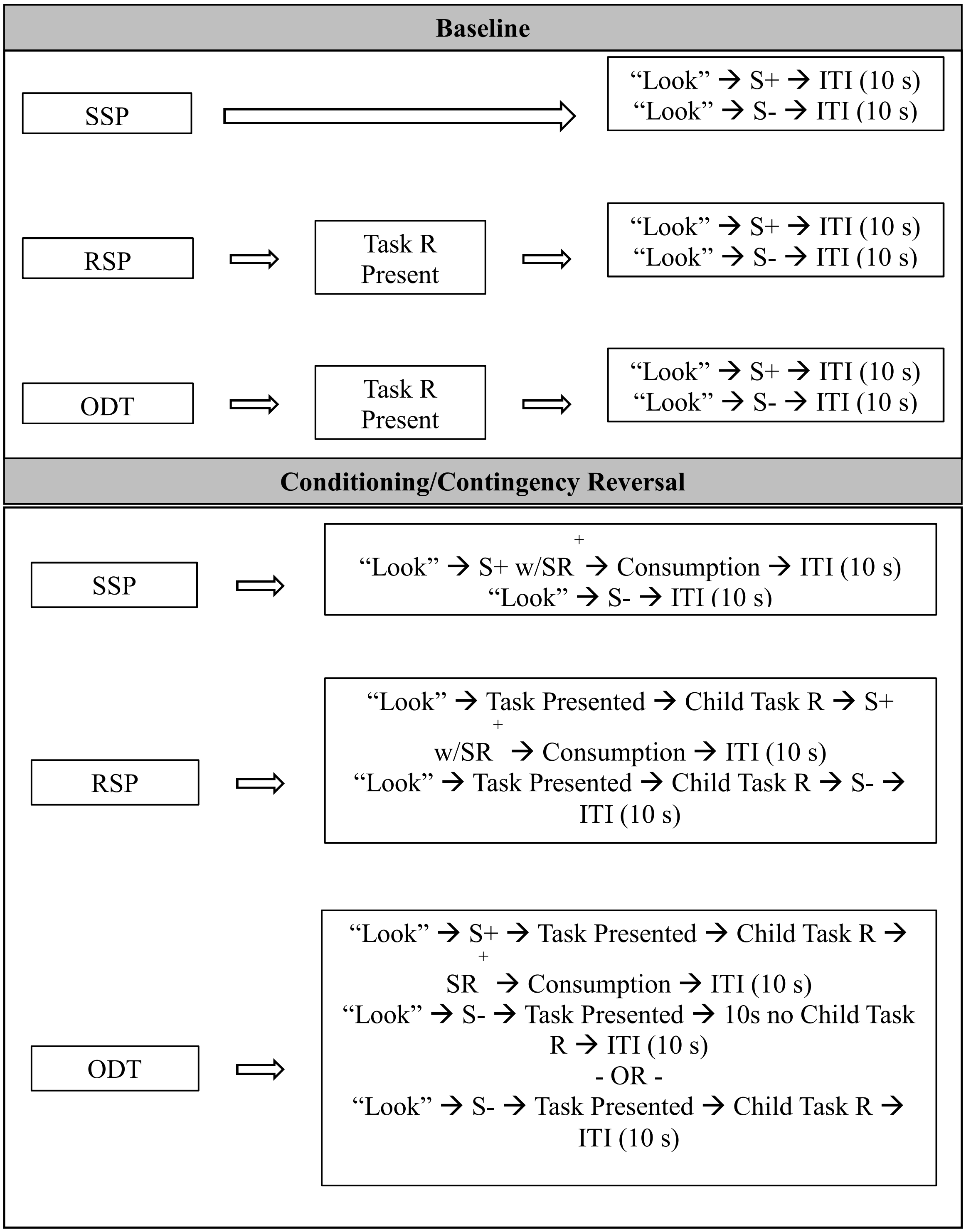
☐ ah	☐ bye bye	☐ one	☐ moo	☐ we
☐ wow	☐ hop	☐ my	☐ up	☐ boy
☐ bee	☐ mama	☐ boo	☐ may	☐ wa wa
☐ knee	☐ papa	☐ no no	☐ pop	☐ toy
☐ oo	☐ me	☐ oh	☐ too	☐ baa

Sub-total Group 1

Assessment			
1st	2nd	3rd	4th

Appendix B

Study 1 Condition Descriptions:



Appendix C

Study 2: Identification of approximation charts

TABLE-4-2 CLASSIFICATION OF THE CONSONANT PHONEMES OF AMERICAN ENGLISH

	PLACE OF ARTICULATION						
Manner of Articulation	BILABIAL	LABIODENTAL	INTERDENTAL	ALVEOLAR	PALATAL	VELAR	GLOTTAL
Stop (oral)							
voiceless	p			t		k	ʔ
voiced	b			d		g	
Nasal (stop)	m			n		ŋ	
Fricative							
voiceless		f	θ	s	ʃ		
voiced		v	ð	z	ʒ		
Affricate							
voiceless					tʃ		
voiced					dʒ		
Glide							
voiceless							h
voiced					j	w	
Liquid				l, ɹ			

	Tongue Position						
Lip Rounding	High-Front Vowels	Mid-Front Vowels	Low-Front Vowel	Mid-Central Vowels	High-Back Vowels	Mid-Back Vowels	Low Back Vowel
Rounded				/ɝ/	/u/ /ʊ/	/o/ /Ɔ/	
Unrounded	/i/ /I/	/e/ /Ɛ/	/æ/	/ɚ/ /ə/ /ʌ/			/ɑ/