October 2019

Varying Inter-Stimulus and Inter-Trial Intervals During Stimulus-Stimulus Pairing: A Translational Extension of Autoshaping

Patricia Eberhardt
peberhardt@rollins.edu

Varying Inter-Stimulus and Inter-Trial Intervals During Stimulus-Stimulus Pairing: A Translational Extension of Autoshaping

A Thesis
By

Patricia Eberhardt

Submitted to the Faculty of the Department of Health Professions
at Rollins College in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF ARTS IN APPLIED BEHAVIOR ANALYSIS AND CLINICAL SCIENCE

July 2019
Winter Park, FL

Table of Contents

Page

ABSTRACT …………………………………………………………………………………4

INTRODUCTION ……………………………………………………………………...5

The Role of Automatic Reinforcement ……………………………………………...6

An Analysis of SSP Based on Autoshaping……………………….................................8

REVIEW OF LITERATURE…………………………………………………..............9

Analysis of Factors Influencing SSP……………………………………………………18

Number of Sounds per Presentation per Pairing and CS Duration………………………...19

Type of Pairing Procedure and Reinforcing Stimulus ……………………………......21

Control Procedures ……………………………………………………………………..22

Relative Temporal Parameters in SSP…………………………………………………..24

STATEMENT OF THE PROBLEM ……………………………………………………..25

METHOD …………………………………………………………………………………28

Subjects and Verbal Responding Levels ……………………………………….....28

Settings and Materials ……………………………………………………………......32

Response Measurement, Interobserver Agreement, and Procedural Integrity………......33

Design and Procedure ………………………………………………………………...34

RESULTS ………………………………………………………………………... 37

DISCUSSION ……………………………………………………………………………..39

REFERENCES ……………………………………………………………………………44

TABLES……………………………………………………………………………...52

FIGURES …………………………………………………………………………………59

Abstract

Stimulus-stimulus pairing (SSP) is a respondent conditioning procedure often implemented to elicit vocalizations in children with language delays. Unfortunately, the research showing the effect of increased rates of vocalizations is mixed. Through analogies drawn between SSP and autoshaping, da Silva and Williams (under review) identified variables potentially responsible for increasing the efficacy of SSP (da Silva & Williams, under review). The present study sought to evaluate the relative duration of the inter-trial interval (ITI) and the inter-stimulus interval (ISI). Specifically, the duration of the ITI was systematically varied from 15 s to 60 s and the value of ISI was proportional to the value of ITI. Nine typically developing children, aged 15 to 21 months participated and were randomly assigned to one of three groups. Pairing (trials in which the sound model preceded the delivery of food) and control (trials in which there was no programmed pairing of the delivery of the sound and food) conditions alternated for all subjects. Results were higher rates of vocalizations in the pairing conditions across all subjects. The ratio of approach/withdrawal to the sound differed systematically among the groups with more approach behavior observed with longer ITIs. Moderate and stable rates of vocalizations were observed in 30-s ITI and 60-s ITI conditions. Contrarily, high but decreasing levels of vocalizations were observed in 15-s ITI condition, with more withdrawal behavior from the CS.

Keywords: Stimulus-Stimulus Pairing; Autoshaping; Respondent Conditioning; Classical Conditioning; Automatic Reinforcement

Introduction

Language is a behavioral cusp and the stepping stone upon which other developmental skills are built (Rosales-Ruiz & Baer, 1997). Young children often exhibit language and vocal communication skills by the first year of their lives (Cherry, 2018). The emerging words and language allow the child to access a variety of reinforcers and environmental events that shape their verbal behavior (Koegel, Koegel, & Suratt, 1992). For instance, a young child can learn words that relate to preferred toys, foods, and people in his or her immediate environment, and then he or she can ask for these items or people effectively and develop social interactions with other individuals. According to the Children’s Hospital of Orange County (CHOC, 2018), typically developing children can imitate animal sounds and use four to six simple words by the first year. By 18 to 24 months, most children begin to use simple phrases or two-word sentences (i.e., “Mommy up”). By 2 years, toddlers may say 100-300 words and can begin to put three words together (i.e., “Me want doll”). By 3 years, children may say 500-900 words and can put four to five words into a sentence. By age 4 (preschool age), children may use four to five words into a sentence and will ask questions frequently. But there is a significant problem when preschool age children fall behind in meeting these milestones upon entering school (Petursdottir & Mellor, 2017). Tomblin et al. (1997) reported that as many as 7.4% of the children in the United States exhibit some specific language impairment (SLI) by the age of 4, which is usually correlated with delays in other areas of the child’s life such as academic progress, social behavior, and adaptive functioning (Tomblin et al., 1997; Petursdottir & Mellor, 2017).

Moreover, children with SLI are at imminent risk for difficulties in reading and certain behavior problems (Tomblin et al., 1997). There is another group of children with neurodevelopmental disorders comprised of children diagnosed with autism spectrum disorder (ASD) that represents 1 of every 59 children in the United States (National Autism Association, 2018). These children typically show deficits in the area of pragmatic language, which is the social use of language, and more often present additional impairments characterized by inappropriate development of vocabulary (Petursdottir & Mellor, 2017). For this population, nearly 40% of individuals with a diagnosis of autism fail to develop vocal communication skills and remain nonvocal for the rest of their lives (National Autism Association, 2018). Thus, effective early language and communication skills interventions are essential to prevent damaging consequences for these children with or without developmental disabilities and with certain language impairment (Sundberg & Partington, 1998; Sundberg, 2008).

The Role of Automatic Reinforcement

Skinner (1957) proposed a function-based analysis of language development, according to which language is learned behavior under the functional control of environmental contingencies like any other behavior (Sundberg & Michael, 2011). Language develops due to the outcome of social and automatic reinforcement (Shillingsburg, Hollander, Yosick, Bowen, & Muskat, 2015). Parents respond socially to most of the infant’s vocal responses (Gros-Louis, West, Goldstein, & King, 2006), and babbling increases in rate and similarity to the intelligibly speech of the parents in part due to the attention the children receive from their parents (Goldstein, King & West, 2003; Wu & Gros-Louis, 2016). Equally important is the role of automatic reinforcement in the acquisition of early vocal responses (Skinner, 1957; Smith, Michael, & Sundberg, 1996). Automatic reinforcement refers to reinforcement that is not mediated by another person (as would be in the case of social reinforcement) but is the direct effect of behavior on the individual's own body or surrounding world (Vaughan & Michael, 1982). The infant's vocal behavior is strengthened by automatic reinforcement when the child is able to reproduce the sounds he hears in the environment (e.g., the sounds of airplanes, cars, vacuum cleaners, etc.) and, most importantly, when the child can match the vocal behavior of his parents (Skinner, 1957; Smith et al., 1996). This process occurs in two stages: pairing and automatic reinforcement. First, the sounds and words that parents emit as they care for their children might be established as conditioned reinforcers through the pairing of their speech with unconditioned reinforcers (e.g., food, warmth, removal of wet diaper, and physical touch) during caregiving routines (e.g., feeding, bathing, diaper changing, and rocking the child to sleep; McLaughlin, 2010; Sundberg & Michael, 2011; Shillingsburg et al., 2015). These same sounds, when produced by the infant during babbling, present the occasion to strengthen the oral muscle movements necessary to produce them (Shillingsburg et al., 2015; Smith et al., 1996). Repeated vocal attempts to match heard sounds and words increase the variety of produced sounds and prepare the infant to speak in words and sentences as a result of automatic reinforcement upon vocal attempts (Carbone, 2012).

The repeated pairing of a sound or word with an established reinforcer serves two purposes (Sundberg, Michael, Partington, & Sundberg, 1996). First, it will make the child's vocalizations be more sensitive to automatic reinforcement when the child hears himself or herself produce a sound or word that share similar acoustic features with the paired sound or word (Petursdottir & Mellor, 2017; Shillingsburg et al., 2015; Smith et al., 1996). The closer the sound production is to matching the sounds that have been conditioned as reinforcers the stronger the reinforcement (Smith et al., 1996; Palmer, 2018). It also creates an opportunity for the behavior analysis practitioner to begin programmed reinforcement procedures to bring the child's emerging sound or word under appropriate stimulus control as functional vocalizations (e.g., echoic, tact, or mand; Shillingsburg et al., 2015).

Researchers have applied this analysis in recent years to the use of a clinical procedure called *stimulus-stimulus pairing* (SSP) in studying development of vocalizations in language-delayed children. Early attempts involved an alternative procedure known in the basic research as *autoshaping* (Myers, 1980). Both SSP and autoshaping rely on respondent and operant processes (da Silva & Williams, under review). First, autoshaping and SSP can generate respondent behavior through response-independent presentations of a neutral stimulus (e.g., sound or word) and unconditioned stimuli (US). Second, the respondent behavior can then be captured by operant consequences (Skinner, 1957; Smith et al. 1996).

An Analysis of SSP Based on Autoshaping

Stimulus-stimulus pairing (SSP) is a procedure in which a neutral stimulus (NS), typically an adult-produced sound or word, is presented or paired with a known preferred item (US: food or CS: tickles, smiles, favorite toy) until NS becomes a conditioned stimulus (CS), such as a spoken sound or word (Smith et al., 1996). Notably, the individual is never required to emit any response (Sundberg et al, 1996). Similarly, autoshaping, which is a procedure commonly used to train pigeons to peck a key in an operant chamber, consists of presenting a NS (e.g., a key light) that predicts the delivery of a US (e.g., grain) and serves to elicit behavior (e.g., keypecking; Brown & Jenkins, 1968; Wilcove & Miller, 1974). Autoshaping, then, involves a stimulus-stimulus contingency (keylight = CS; grain = US) that is similar to the one in SSP, where the presentation of a CS (e.g., adult's vocalization) and a US (e.g., food, toy, coin) might result in CR during or following the CS presentations. This compatibility allows for further direction in the study of the emergence of language for nonvocal children, where early vocalizations might be acquired by autoshaping/SSP procedures. The value of identifying common procedural features between autoshaping and SSP is the opportunity to resolve some limitations of SSP, mainly that it has mixed effectiveness (Militios et al., 2012). Many researchers have attempted to identify variables that can be called to be responsible for the procedure's effectiveness or failure, which has led to several distinct procedural variations of SSP. To date, researchers have not been able to make recommendations to behavior analysis practitioners based on reliable results (Petursdottir & Lepper, 2015). Thus, successful applications in autoshaping should be extended to SSP.

Review of Literature

Autoshaping has been reported to occur with different species of nonhuman subjects and using different types of responses and reinforcers (Jenkins & Moore, 1973; Sidman & Fletcher, 1968). Brown and Jenkins (1968) demonstrated that autoshaping can produce behavior without manually shaping such response. The authors found that pigeons would peck a light up key even though the grain presentation was independent of the response. Gamzu and Schwam (1974) applied autoshaping to the key pressing response of squirrel monkeys and found similar results. Some studies have extended autoshaping procedures to human subjects. In particular, Myers (1981) evaluated the effect of autoshaping in human's vocal responses. In Experiment 1, three typically developing infants (16 to 18 months) participated. The subjects stayed in their cribs or playpen during the training. The experimenter vocally presented a sound "Q" (NS) once and immediately delivered a small bite of food (US; ice cream, sherbet, banana, etc.). A fixed intertrial interval (ITI) began after 5 s of food consumption and the subjects were allowed to play with their toys. The ITI for Subject 1 and Subject 3 was 60 s throughout the study. The ITI for

Subject 2 was 30 s for nine sessions and 60 s for the last (10th) session due to early termination. All three subjects exhibited an increase in target vocalization. The experimenter terminated the training once subjects met a criterion of 15 successive trials with a CR. The control condition was similar to the autoshaping condition, but the CS was no longer followed by the US. The condition was terminated once the subjects met criterion of a minimum of five successive trials without a CR. An autoshaping condition followed control. The subjects were re-exposed to the conditions as in the first autoshaping procedure. Subjects met the autoshaping criterion during the second session of this condition and they were able to produce target sound.

The use of autoshaping to establish infant vocalizations was short lived, but the interest in establishing vocalizations in children was revived by Sundberg et al. (1996). These authors conducted one of the first empirical applications of SSP in the examination of the function of automatic reinforcement in the acquisition of vocal responding in children with language delays. The study included five children with ages ranging from 2 to 4 years with a wide spectrum of language abilities. One subject was a typically developing child who showed age-appropriate language skills. Four subjects had diagnoses of moderate to severe language delays. These four subjects were able to engage in more than 100 mands, tacts, and intraverbals, but only three of them exhibited some vocalizations, whether vocal speech (i.e., words) or vocal play (i.e., sounds). The study had three conditions, pre-pairing (baseline), pairing, and post-pairing. In the pre-pairing condition, the researchers recorded all the subjects' vocal responses emitted during free play as well as the controlling variables observed (e.g., an establishing operation that was in effect or exposure to nonvocal or verbal discriminative stimuli). The experimenters did not interact with the subjects during this condition. As a result of this condition, the authors selected sounds, words, or phrases known to be novel for each child as their target responses that later were introduced in the pairing condition. In the following condition, a familiar adult conducted the pairing procedure by presenting the target vocal, varying the pitch and intonation on every presentation, and immediately following the vocal by a form of social interaction (e.g., tickles, praise, spinning, clapping, bouncing, or animated parental attention). The children were never required or prompted to repeat the sound, word, or phrase. The experimenters conducted, on average, 15 pairings per minute across subjects and sessions. Repeated pairings were intended to elicit vocalizations by the children that duplicated the adult's vocalizations and the target vocals did not topographically relate to any of the paired reinforcing stimuli (e.g., the target word "rock" could be paired with tickles). After the pairing condition, the adult stepped back and recorded data as in pre-pairing condition. Results showed that all five subjects acquired novel vocal responses after the implementation of the stimulus pairing procedure and these responses were exhibited in the post-pairing sessions. Moreover, the authors also noted that there was a significant increase in overall vocalizations rate across subjects, but the effects were temporary. The implications of the study were important to open another possibility to treat language delays in children without the application of reinforcement via shaping, echoic training, or prompting procedures (Yoon & Bennett, 2000).

One could claim that the vocal responses acquired by Sundberg et al.'s (1996) subjects were a result of the simple exposure to the vocalizations or by accidental contiguity between child's vocal responses and the delivery of preferred social interactions (Palmer, 2018). To the contrary, the relevant variable was the temporal association between the adult-vocal model and strong reinforcers (Petursdottir & Lepper, 2015). That is, the adult-produced sounds elicited similar sounds by the children, ostensibly through the process of respondent conditioning (i.e., *antecedent pairing*; Sundberg et al., 1996). These results aligned with Bloom and Esposito's

(1975) early findings that proposed social stimulation can function as either an eliciting (as in respondent conditioning) or a discriminative (as in operant conditioning) stimulus that precedes the infant's initial vocalizations rather than functioning as a reinforcer alone. Likewise, early vocalizations might be respondent behavior (e.g., crying and screaming) that become operant responses if such vocalizations are followed by reinforcement (Sundberg et al., 1996). Thus, initial vocalizations might be elicited by conditioned stimuli (CS) or unconditioned stimuli (US) and evoked by discriminative stimuli (SD) or establishing operations (EO; Bloom & Esposito, 1975).

Smith et al. (1996) extended the previous study by examining SSP under neutral, negative, and positive conditions to further demonstrate that the simple exposure to the target sound was not sufficient to produce an effect. For this purpose, the authors arranged a neutral condition in which the experimenter presented a target vocal response but did not deliver any reinforcing stimulus. Furthermore, the added negative condition was intended to demonstrate that the pairing procedure for establishing a NS as an automatic punisher was the same as the procedure for establishing a NS as an automatic reinforcer, with the exception that the NS was paired with a form of punishment (e.g., reprimand). In the positive (pairing) condition, the target vocal response was followed by a preferred item (e.g., bubbles or tickles). As in Sundberg et al.'s (1996) study, the experiment included the same three phases (e.g., pre-pairing, pairing, and post-pairing) in which the experimenters recorded targeted and non-targeted vocalizations. Two typically developing female subjects (11 and 14 months of age, respectively) participated in the study. The target vocal response was defined as emission of recognizable phonemes that were observed previously in the subjects' repertoires at low rates or were not observed during pre-pairing, and the vocal stimulus was presented only once per pairing. Subject 1 participated in all three pairing conditions and Subject 2 only participated in the positive condition. Results demonstrated that the neutral pairing produced no change in the rate of production of target vocal responses in post-pairing condition, the positive condition produced a distinct increase in the rate of target responses as well as a lesser increase of the rate of non-target vocalizations for both subjects and, finally, in the negative pairing the target vocal immediately stopped and overall rate of responding decreased too. The authors concluded that both automatic reinforcement and automatic punishment influence differently the acquisition and development of infant's vocal verbal behavior. This finding also highlights the significant function of the reinforcing stimulus during the pairing of sounds with reinforcers to increase the rate of vocal responding (Petursdottir, Carp, Matthies, & Esch, 2011).

Yoon and Bennett (2000) conducted two experiments to investigate the role of pairing procedures and automatic reinforcement in the acquisition of vocal responses and extended SSP to children with developmental and language delays. In Experiment 1, the authors replicated Sundberg et al.'s (1996) pairing procedures to establish new vocal sounds in the repertoire of three preschool age children (3 and 4 years old) who showed zero levels of vocal response and limited vocal play and listener responding skills. The study used a multiple-baseline-across-subjects design with the same three conditions as in previous research: pre-pairing, pairing, and post-pairing. A specific sound was selected for each subject and paired with physical interaction. After a pre-pairing condition was conducted as in previous studies, a pairing session was initiated with approximately 12 pairings per minute over 3 min session. A post-pairing session immediately followed the pairing condition, which was similar to the pre-pairing procedure. Results corroborated findings from previous studies (Smith et al., 1996; Sundberg et al., 1996), and target vocal responses occurred for all subjects immediately after the pairing procedure.

Likewise, the authors noted that the effects of pairing were temporary with a noticeable decline in frequency of vocalizations after approximately 9 min of stopping the pairing. In Experiment 2, Yoon and Bennett (2000) compared SSP with echoic training to verify if the results from Experiment 1 were due to stimulus pairing or operant reinforcement. Two subjects from the previous experiment and a new subject with similar verbal behavior as the other two subjects participated in the experiment. The study included pre-echoic, echoic, post-echoic, pairing, and post-pairing conditions. A multiple-baseline-across-subjects design was used to demonstrate experimental control. The selected target vocal sounds were not observed during baseline condition. The pairing procedure was identical to Experiment 1 and was used to establish target vocal sounds. In echoic training, the experimenter modeled a vocal sound for the child to vocally imitate and delivered reinforcement contingent on the echoic response. The results in the pairing and post-pairing conditions were similar to those in Experiment 1, which indicated an initial increase in target vocal sounds at the begin of the post-pairing session and decreasing levels of responding after a certain period of time. The results in the echoic training suggested that the echoic condition had no immediate effect on target vocal responses. Therefore, the authors found that SSP was more effective than echoic training at increasing the vocalizations of four preschoolers with severe language and communication delays. This study was thus able to show that even for individuals with severe language and developmental delays, vocal responses can be strengthened by automatic reinforcement when stimulus pairing procedures are utilized.

In spite of these promising findings, it has been argued that in these previous studies (Sundberg et al., 1996; Smith et al., 1996; Yoon & Bennett, 2000), appropriate experimental control for elicitation of vocalizations in the demonstration of automatic reinforcement required more robust methodology (Petursdottir, Carp, Matthies, & Esch, 2011). When this limitation is accounted for, however, there remains support for SSP. For example, Esch, Carr, and Grow (2009) used a modified SSP procedure that consisted of interspersing paired and unpaired trials, with the latter procedure that serving as a control. The unpaired stimulus (S-) did not predict the delivery of reinforcers, unlike the paired stimulus (S+) which was followed immediately by a reinforcing stimulus. An observing prompt (e.g., “Look”) was used to redirect the children’s attention toward the experimenter before initiating any trials. “Motherese” modeling, similar when mothers engage in “baby talk” with their children, was used across all trials to enhance the stimulus salience during the pairing. Three subjects (aged between 2 and 5 years) with a diagnosis of autism were selected for the study. The subjects represented a wide range of vocal and verbal behavior abilities. Subject 1 had no echoic responses and low frequency vocal play, Subject 2 engaged in frequent vocal play without any functional verbal behavior, and the third subject displayed frequent vocal play but had few mands, tacts, and intraverbals. The study had four conditions. The baseline, SSP, and programmed reinforcement conditions were conducted within every trial. A noncontingent reinforcement (NCR) condition was added and consisted of delivering reinforcers for 5 min on a fixed-time (FT) 30-s schedule. Any target vocal responses that occurred within 5 s of scheduled reinforcement were followed by a 20-s correction delay to avoid adventitious reinforcement of responding. Results indicated that target vocalizations increased during SSP to satisfactory but moderate levels over baseline for all three subjects.

Rader et al. (2014) extended Esch et al.’s (2009) study to show the generality of the improved SPP methodology. Three subjects (4, 6, and 7 years of age) diagnosed with autism and exhibiting low levels of vocal play participated. Two of the three subjects showed remarkable increases in the rate of vocal responses for the target vocalization after the pairing procedure but showed minimal increases for the non-target vocal response, which confirmed the effectiveness of SSP. The authors attributed the failure of SSP to increase vocalizations for the third subject to the child's problem behavior that might have interfered with the training rather than being due to the SSP procedure itself. Barry, Holloway, and Gunning (2019) sought to replicate and extend Esch et al.'s (2009) findings to the application of parents as pairing agents. The study included two subjects (2 and 4 years old) diagnosed with autism who had no echoic or vocal repertoires. The subjects' parents conducted five conditions: baseline, SSP, direct reinforcement, NCR, and direct reinforcement again. The study included a parent training on SSP before the experimental condition. The experimenter visually prompted the parents delivering either the target or non-target sounds across conditions. The authors also added a randomly-determined ITI that ranged from 5-30 s; sounds were presented once per trial. Results suggested that SSP can be implemented successfully by parents, as both children demonstrated higher levels of vocal responding across all conditions when compared against baseline. These studies (Esch et al., 2009; Rader et al., 2014; Barry et al., 2019) provide support of the effectiveness of SSP. Indeed, these findings showed that SSP can be an adequate procedure to increase vocal responses that in other way have not shown to be sensitive to reinforcement.

Nevertheless, there are other studies in which SSP was effective for some subjects and ineffective for others (Carroll & Klatt, 2008; Miguel, Carr, & Michael, 2002; Stock, Schulze, & Mirenda, 2008; Yoon & Feliciano, 2007). These findings have challenged the application of SSP in clinical settings and, as a result, researchers have called for more research to determine clinical significance of SSP (Petursdottir & Lepper, 2015). Several distinct variations of SSP have made it difficult to identify factors influencing successful conditioning of target sounds. Some main differences have been identified as the number of sound presentations, type of pairing (e.g., delay, trace, etc.), number of trials per pairing, and ITI duration, among other factors (Shillingsburg et al., 2015).

Petursdottir et al. (2011) evaluated certain variables (e.g., frequency of preference assessments, elimination of pre-session exposure to target sounds, number of pairings, use of observing prompt, alternating pairing trials with control trials, and using pre-recorded sounds) that were believed to play a role in the efficacy of SSP. The study included three boys diagnosed with autism. Two subjects were 4-year-old twin brothers who had high levels of vocal play, echoic, and mands. The third subject was a 3-year-old boy who had limited vocal play and no functional verbal behavior responses. The subjects were taught to press a button from a two-button device that resulted in the immediate production of either the target or control sound. The allocation of the responses and sounds produced by the subjects were recorded across baseline and pairing sessions in two experimental phases. In Phase 1, the experimenters presented target sounds via a computer in randomized blocks of three and with a 10-s ITI (buttons and reinforcers were out of the subject's reach during baseline sessions). In addition, experimenters used an observing prompt to gain subject's attention prior to presenting the stimuli. The target sound and control were presented three times per pairing across 10 trials per target and 10 trials per control sound, along with 10 presentations of the selected reinforcing stimuli. In the pairing sessions, the experimenter used the observing prompt, target sound, and control sound as in the baseline, but only the presentation of the target sound was followed by the reinforcing stimuli. The sessions included 10 pairings for the target sound and 10 pairings for the control. Additional conditions were added for some subjects such as extended pairing, pairing without pre-exposure, and testing. In the extended pairing, Subject 3 was exposed to 20 pairing of the target sound with the reinforcer and 20 presentations of the control sound. In the pairing without pre-exposure, Subjects 2 and 3 were exposed to two novel sounds in a procedure similar to extended pairing. Testing included 2-min free access to the buttons following each sound presentation. Subjects were physically prompted to press the buttons as the experimenter held the buttons in each hand. In Phase 2, only Subjects 2 and 3 participated. In the baseline, buttons were arranged as in the test condition and the experimenter prompted the subjects to press the button every 20 s. Similar procedures were followed in a continuous (i.e., Fixed Ratio 1) schedule of reinforcement, but prompted and unprompted button-pressing responses were followed by an edible reinforcer or toy. This condition was followed by an extinction phase, which was identical to baseline. Between FR 1 and extinction, the location of the target sound was alternated. Results showed that SSP effects on the selection of target sound and allocation of the target sound to the corresponding button were only evident in Subject 1. The findings did not offer a better interpretation of the effect of manipulated variables that could reliably increase the effectiveness of SSP.

Analysis of Factors Influencing SSP

A closer look at some of the SSP research that attempted to replicate or extend previous findings may allow reviewers to identify common procedural characteristics. For example, Carroll and Klatt (2008), Miguel et al. (2002), and Stock et al. (2008) used a modified SSP procedure that consisted of a delay conditioning with 5 presentations of target sounds per pairing and a fixed 0-s (Carroll & Klatt, 2008; Miguel et al., 2002) or 20-s ITI (Stock et al., 2008). The authors also arranged 20 (Carroll & Klatt, 2008; Miguel et al., 2002) and 30 (Stock et al., 2008) trials per session and selected edible and tangible reinforcers. As a result, four of eight participants responded positively to SSP. These procedural modifications differed substantially from the procedures used by Barry et al. (2019), Esch et al. (2009), Rader et al. (2014), in which all 10 of 10 subjects benefited from SSP. For instance, the latter authors used trace conditioning with variable ITI durations (e.g., 5-30 s duration), three sound presentations per pairing across 20 trials per session, and also both edible and tangible reinforcers. Then there is another group of studies (Yoon & Bennett, 2000; Yoon & Feliciano, 2007) that employed a simultaneous pairing procedure with one sound presentation per trial across 36 trials per session and with social, edible and/or tangible reinforcers. In this case, six of nine participants benefited from SSP procedure.

If one could analyze these variables as in isolated clusters of research instead of among all and across all of the studies, and under a respondent conditioning view, one may find new information that might lead researchers to a more refined SSP methodology. The following variables are analyzed under the line of autoshaping research as suggested by da Silva & Williams (under review).

Number of sound presentations per pairing and CS duration. Researchers have varied the number of sound-presentations in SSP from one presentation (i.e., Sundberg et al. 1996; Ward, Osnes, & Partington, 2007; Yoon & Bennett, 2000; Yoon & Feliciano, 2007), three presentations (Esch, Carr, & Michael., 2005; Esch et al., 2009; Lepper, Petursdottir, & Esch, 2013; Rader et al., 2014), five presentations (Carroll & Klatt, 2008; Miguel et al., 2002; Stock et al., 2008), and seven presentations (Normand and Knoll, 2006). Miliotis et al. (2012) systematically manipulated the number of sound presentations (e.g., one versus 3 sounds per pairing) as a possible variable that influence SSP effectiveness. Two subjects (6 and 8 years old) diagnosed with autism participated in this study. Subject 1 had low levels of vocal play and no echoic responses. Subject 2 also had low levels of vocal play but exhibited six echoic responses. The study included four conditions where the presentation of a target sound (S+) and nontarget sound (S-) were varied from one to three repetitions per pairing trials. Both subjects exhibited significantly increased levels of vocalizations when the target sound was presented only once per pairing trial (as reported by Smith et al., 1996; Sundberg et al., 1996; Yoon & Bennett, 2000; Yoon & Feliciano, 2007). Yet, there is evidence that SSP can work with different variations of sound presentations (e.g., 3 or 5 sound-presentations).

Shillingsburg et al. (2015) suggested that there did not appear to be a relationship between the number of sound presentations per pairing and the magnitude of SSP effectiveness after conducting a nonoverlap of all pairs analysis (NAP), which it can tentatively lead to the conclusion that the number of sound presentations per pairing might not be a relevant variable influencing the effectiveness of SSP. However, if the duration of the number of sound presentations is measured, one could identify different durations of CS per pairing trials based on the length of the sound (e.g., banana or bah) and the number of times it is repeated that may provide different information. For instance, in autoshaping, there seems to be a relationship between the duration of CS and autoshaping effectiveness (Brown & Jenkins, 1968). According to Ricci (1973), when 30-s key lights preceded food, pigeons achieved faster autoshaping than when 120-s key lights were presented. Conversely, Brown and Jenkins (1968) found that 8-s key light presentations were more effective than 3-s key light presentations. da Silva and Williams (under review) explained that there appears to be an optimal length of CS that influences autoshaping, and that the optimal CS duration depends in part on the ITI duration. Thus, the success or failure of SPP might depend partly on the relationship between the duration of the sound presentations and ITI duration. For example, one could explain that the failure of establishing vocalizations in Esch et al.'s (2005) study was due to an inferior relationship between ITI 0-s and three experimenter-sound productions per pairing, among other possible variables. Therefore, in SPP, it probably is more valuable to 1) measure the length of the sound presentations (e.g., duration of CS in seconds) than counting the number of sound presentations and 2) determine an optimal duration of CS and ITI as variables that impact SSP. Using a pre-recorded sound that represents a selected duration could facilitate consistent CS presentations across pairings and in relation to the ITI duration.

Type of pairing procedure and reinforcing stimulus. As mentioned above, some of these studies have used delay conditioning in which the presentation of the reinforcing stimulus (US) occurs while the sound (CS) is still present (e.g., Carroll & Klatt, 2008; Miguel et al., 2002; Lepper et al., 2013; Miliotis et al., 2012; Normand & Knoll, 2006); trace conditioning where presentation of US occurs at the point of or after the conclusion of the CS (e.g., Barry et al., 2019; Esch et al., 2009; Smith et al., 1996; Sundberg et al., 1996; Rader et al., 2014); or simultaneous conditioning in which the onset of US occurs at the same time of CS onset (e.g., Yoon & Bennett, 2000; Yoon & Feliciano, 2007).

Lepper et al. (2013) have been the only ones to conduct an SSP study to evaluate delay and discrimination pairing procedures. The authors reported that both pairings were similarly effective. Moreover, Shillingsburg et al. (2015) reported that delay conditioning drew stronger size effects when it was employed over trace and simultaneous conditioning. This result aligns with Brown and Jenkins (1968) affirmation that forward pairing, in this case like delay pairing, is more effective to conditioning target response than simultaneous or backward pairing. In addition, trace conditioning – another type of forward pairing - has also shown to be effective in autoshaping (e.g., Brown & Jenkins, 1968) and SSP (e.g., Barry et al., 2019; Esch et al., 2009; Rader et al., 2014; Smith et al., 1996; Sundberg et al., 1996). In contrast, two of the SSP studies that used simultaneous pairing have achieved mixed results. These results could be explained by the possible relationship between the type of reinforcing stimulus and type of pairing procedure used. da Silva and Williams (under review) explained that if simultaneous conditioning must be used, it is recommended to pair sounds with social reinforcers. The authors explained that social reinforcers are least likely to interfere with the target vocal response when presented at the same time. Evidence in support of this recommendation was observed in Yoon and Bennett's (2000) study, in which authors successfully paired vocalizations with social reinforcers for all three subjects. Hence, in Yoon and Feliciano's (2007) study, the authors employed simultaneous pairing but used social interactions, edibles, and tangible items as reinforcers. Only three of six subjects showed increase levels of vocalizations over baseline. It could be possible that, for those subjects who received edible reinforcers, chewing the food might have interfered with the subject's production of the target sound as both sound and edible reinforcer stimuli occurred at the same time (da Silva & Williams, under review). Therefore, it appears to be a good practice to use trace or delay pairing if edible/tangible reinforcers are selected because the presentation of the stimuli are not occurring at the same time and would not interfere the display of target response (da Silva & Williams, under review). Moreover, in autoshaping research, trace conditioning and edible reinforcers have been extensively used and strong results have been obtained (Brown & Jenkins, 1968; Kaplan & Hearst, 1982; Kaplan, 1984; Lucas, Deich, & Wasserman, 1981).

Control procedures. Shillinsgburg et al. (2015) reported that 8 of 13 SSP studies used similar control procedures to prevent adventitious reinforcement. In general, this means the researchers introduced a brief delay (e.g., 20 s, Militios et al., 2012; 30 s, Normand & Knoll, 2006) before delivering the reinforcing stimulus (US) when the subjects exhibited the target response (CR) following the presentation of the auditory stimulus (CS; Carroll & Klatt, 2008;

Esch et al., 2009; Militios et al., 2012; Rader et al., 2014). As a result, half of these showed mixed results (Carroll & Klatt, 2008; Miguel et al., 2002; Stock et al., 2008) and failed to produce vocalizations (Normand & Knoll, 2006).

In autoshaping, this type of control procedure is similar to *omission training* (Williams & Williams, 1969), a procedure in which an edible reinforcing stimulus occurs only on those trials in which CR (e.g., keypecking in pigeons) did not occur. Locurto, Terrace, and Gibbon (1976) explained that omission training is relevant to evaluate stimulus-stimulus contingencies (S-S) in autoshaping, and one could argue similarly in SSP, because there is a greater chance that S-S pairings close in time can establish a direct contingency on the CR. The point of this control procedure when applied in SSP research is to reduce the likelihood of operant contingencies controlling verbal behavior to better isolate and understand the control of behavior by respondent conditioning (Miguel et al., 2002). Rescorla (1967), however, explained that this control procedure not only removes the contingency between CS and US but in fact adds a new contingency, which it is that the US cannot follow the CS for the duration of the delay. As a result, the CS comes to serve as a signal for the absence of the US instead of the signal for the presence of the US (the exact opposite of the learning that should be facilitated in SSP). With this in mind, it is possible that Normand and Knoll’s (2006) subjects failed to reproduce target sounds because the use of a delay control procedure successfully paired the withholding of the reinforcing stimulus when subjects exhibited the target vocalization following CS (da Silva & Williams, under review) and not due to SSP inefficacy.

In autoshaping of nonhuman responses, omission training can retard the re-establishment of CR after subjects are switched back to pairing conditions (Experiment 1, Locurto et al., 1976). This inhibitory effect could have been exacerbated in part due to the limited number of training trials provided (e.g., 10 trials per condition) in this study, which could have impeded the appropriate pairing of stimuli after switching from control to pairing condition. Rescorla proposed an alternative control procedure, a *truly random control* procedure, based on the logical framework in which CS and US are presented on independent random schedules without any contingency between the stimuli. Rescorla suggested that this control procedure seemed more adequate than other conventional control procedures for respondent conditioning because it is based on the idea that the CS-US contingency is important instead of the CS-US pairing. A relevant feature then is that the occurrence of the CS gives no information about the occurrence of the US; consequently, it eliminates the new contingency established by the omission procedure. Despite Shillingsburg et al.'s (2015) recommendation of using the omission training as a control procedure, the disadvantage of doing so it seems counterproductive for the progress of SSP.

Temporal parameters in SSP. One overlooked variable in SSP is determining the optimal CS-US interval, which is simply the duration between the presentation of the CS (i.e., the sound) and the presentation of the US (i.e., the preferred item, food, or social interaction). In autoshaping, a critical factor in determining the optimal CS-US interval is its relative duration to the time between trials, or ITI. Kaplan (1984) measured approach responses to the CS and withdrawal responses from the CS to evaluate the relative duration of the ISI to the ITIs. In Experiment 1, a keylight (CS) was presented for 12 s, and followed by a 12-s inter-stimulus interval (ISI). The duration of the ITI was systematically varied in duration between 15 s to 240 s across five groups of pigeons. It was expected that pigeons would display different acquisition responses. In fact, excitatory conditioning or approach to the keylight occurred when the ITI was long (e.g., 240 s); inversely, inhibitory conditioning (withdrawal from the keylight) occurred with short ITI (e.g., 15 s). Results indicated that pigeons tended to approach the CS when the ITI was greater than 60 s, and they showed withdrawal responses when the ITI was less than 60 s.

One important outcome of Kaplan's study for SSP research is that the failure to observe CR (i.e., target sound or word) is not explained by a failure of the procedure itself. Instead, failure can be explained due to excitatory and inhibitory conditioning of equal strength cancelling out each other. In short, the application of Kaplan's findings to SSP research can aid in identifying the optimal CS-US length (i.e., trial duration) to conditioning target vocalizations. One major limitation, however, is that the trial duration and ITI duration are usually not reported in SSP studies (e.g., Yoon & Bennett, 2000; Yoon & Feliciano, 2007; Normand & Knoll, 2006).

Statement of the Problem

The implications of Sundberg et al.'s (1996) results were important to demonstrate the role of automatic reinforcement in increasing vocalizations and the effectiveness of SSP. These findings thus opened another possibility to treat language delays in children with or without developmental disabilities (Yoon & Bennett, 2000). Since the publication of Sundberg et al.'s (1996) paper, researchers have investigated the effectiveness of SSP to increase vocalizations of children with language delays. Unfortunately, behavior analysts have yet to reach a consensus on effectiveness of SSP. Sundberg et al. (1996) and others (Carroll & Klatt, 2008; Esch et al., 2009; Miguel et al., 2002; Militios et al., 2012; Smith et al., 1996; Rader et al., 2014; Stock et al., 2008; Yoon & Bennett, 2000; Yoon & Feliciano, 2007) found SSP to be successful for all or some of the participants but Esch et al. (2005) and Normand and Knoll (2006) did not. Shillingsburg et al. (2015) indicated that recommendations are yet to be made because several distinct variations of SSP have found mixed results. However, it might be important to further investigate this procedure for its potential clinical benefits for nonvocal children because, when SSP works, it produces significant clinical gains (Pettursdottir et al., 2011).

Shillingsburg et al. (2015) identified several of the variables that have been included in many of these past studies (e.g., participants' age and diagnosis, type of pairing, number of sound-presentations, type of reinforcer, control procedures) and evaluated their influence on SPP. Unfortunately, the authors found limited conclusions but urged others to continue researching the role of these variables. da Silva and Williams (under review) suggested that the procedural parallels between autoshaping and SSP can answer some of these procedural failures and eventually aid to the establishment of a robust methodology for the conditioning of human vocalizations. Evidence from autoshaping research (see Kaplan, 1984) has shown that manipulations of ITI duration had a significant impact on conditioning. In Kaplan's (1984, Experiment 1) study, pigeons exposed to ITI durations longer than 60 s (e.g., 120-s and 240-s ITI conditions) acquired the approach response to CS. Inversely, pigeons exposed to ITI durations shorter than 60 s (e.g., 15-s and 30-s ITI conditions) exhibited mild withdrawal response from CS. One could predict that, by manipulating the trial duration in relation to the ITI duration as in Kaplan's study, we might obtain relevant information about conditioning human vocalizations. It is possible that studies that obtained positive outcomes for all of the participants (Esch et al., 2009; Lepper et al., 2013; Miliotis et al., 2012; Rader et al., 2014) were able to establish an optimal relative duration between ITI and ISI as they maintained the ISI constant and vary ITI duration (e.g., 5-30 s). In addition, da Silva and Williams (under review) suggested that varying ITI duration might enhance the unpredictability of US presentation, which in turn might result in more vocalizations.

In addition to the short ITI durations that have been reported to date (e.g., 5 s-30 s), longer ITI lengths should be evaluated based on Kaplan's findings discussed previously (e.g., Myers, 1981, used 45 s or 60 s). Human subjects exposed to 15-s and 30-s ITI conditions might exhibit mild to strong withdrawal response from CS. By contrast, subjects exposed to longer ITI durations might demonstrate approach response to the CS. To date, no studies have evaluated the role of the temporal relationship between the ITI and the ISI. We predict that these variables might have a critical impact on the conditioning of human vocalizations as has been observed in other species (e.g., Jenkins & Moore, 1973; Sidman & Fletcher, 1968). It is expected that humans and nonhuman animals would be similarly sensitive to the values of ISI and ITI.

Shillingsburg et al. (2015) reported 75% of participants were young children (e.g., preschool age) and, of these, 69.2% of the participants were diagnosed with autism and 30.6% had other diagnoses (e.g., educational delay, developmental delay). Shillingsburg et al. excluded three of the 42 total participants across studies from their review because they were typically developing children (although they responded positively to SSP). Instead, one should consider extending SSP to typically developing children before proceeding its application to children with language delays or other diagnoses because it is possible that the effects of SSP might differ from those children with language delays and typical developing children (Sundberg et al. 1996). Also, these children are more likely to orient their faces and attend to an adult face and voice (Vouloumanos, Druhen, Hauser, & Huizink, 2009). These responses are considered to be significant *sign-tracking* behaviors. Myers (1981) proposed that sign tracking (i.e., direction) is, in fact, the precursor of the CR. da Silva and Williams (under review) agreed with Myers and explained that one should consider including participants with specific characteristics that will make them more responsive to SSP. This study, therefore, sought to evaluate the temporal relationship between the ITI and the ISI in the establishment of human vocalizations.

Method

Subjects and Verbal Responding Levels

Nine typically developing children between the ages of 15 and 21 months old participated in this study. The subjects were recruited by word-of-mouth and were selected to participate because their parents reported that the children were reaching their developmental milestones and did not appear to have any health or medical concerns.

The Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) and the Early Echoic Skills Assessment (EESA; Esch, 2008) were used to assess subjects' verbal repertoires (specifically mand, tact, listener, and echoic responses). Table 1 displays the VB-MAPP results across subjects. In addition, the experimenter videotaped the subjects during a 30-min free play period in which the frequency and topography of each child's vocalizations were recorded.

Results of the VB-MAPP (Sundberg, 2008), EESA (Esch, 2008), and direct observation for Nolan, who was a 21 months boy, indicated low rates of vocal play (1.27 sound per min) and emitted vowel sounds like "Ah", "Oh," "Uh-Oh," "Ed" for red and consonant-vowel words in the form of tacts like "Horse," "Sheep," "Dog," "Book," "Fog" (for frog), and as a mand such as "Ball" and "Book" (VB-MAPP Level 1, 13.5 of 20 points; EESA Level 2). Assessment results for Ben, a 16-month-old boy, indicated high rates of vocal play (2.07 sounds per min) and emitted vowel sounds like "Asit," "Uh-oh," "Eech" and consonants sounds "See Ya," "Chututu." As per parent report, Ben was able to say "Dada," "Night-night," "Bye-bye," "Shoe" "Dada," "Rusty," and "Mama" but only shoe (2x as a tact) and dada (8x as a form of tact) were observed during the assessment. (VB-MAPP Level 1, 9; EESA Level 1). Assessment results for Ken, who was a 19-month-old boy, showed high rates of vocal play (2.13 sounds per min), consisting of vowels like “Eee,” and one word, “Book,” as a form of echoic-mand (VB-MAPP Level 1, 7 of 20 points; EESA Level 0). Assessment results for Nelson, who was a 21-month-old boy, indicated very low rates of vocal play (0.433 sounds per min) and no intelligible words. His parent expressed that Nelson has never spoken at home (VB-MAPP Level 1, 5 of 20 points; EESA Level 0). Assessment results for Carla, an 18-month-old girl, indicated low rates of vocal play and emitted 10 words including “Mermaid,” “Baby,” “Bottle,” “Please,” and “Star.” Carla was able to reproduce eight animal sounds and tacted fish as “Bloop-bloop” (VB-MAPP Level 1, 18 of 20 points; EESA Level 3). Assessment results for Eva, who was an 18-month-old girl, indicated low rates of vocal play (1.17 sound per min) and vowel sounds like “Uh-oh” (as a form of mand for attention), “Ee-ee,” “Ah,” “Oiaa,” (VB-MAPP Level 1,13 of 20 points; EESA Level 1). Assessment results for Sam, who was a 15 months boy, indicated low rate of vocal play (0.43 sounds per min) including “Ee-ee” and only one word (e.g., What?) (VB-MAPP Level 1, 8 of 20 points; EESA Level 0). Assessment results for Tess, a 17-month-old girl, showed very low rates of vocal play (0.6 sounds per min). She emitted eight animal sounds, tacted “Star,” “Shoe” and “Ook for book, “Oat” for boat, “Keys,” “Kup” for cup, and “Nana” for banana. Tess also showed early mands such as “More,” “Bubbles,” “Wawa” for water, and “Ook” for book (VB-MAPP Level 1, 17 of 20 points; EESA Level 2). Assessment results for Matt, who was an 18-month-old boy, indicated a low level of vocal play (1.06 sounds per min), included eight words including “Kitty cat,” as well as animal sound meow, woof, and roar, and sounds like “Oh!” “Yay”, and “Ouch.” However, his parent reported that Matt can say about 18 words but they do not appear to be under stimulus control (e.g., ball, book, jump, dada, mama, bike, baby, eye, nose, etc.; VB-MAPP Level 1, 13 of 20 points; EESA Level 2).

Selection of target vocal stimulus per subject. The experimenter selected one word unique for each subject based on the following criteria: 1) the word was known to be novel, 2) the word was produced infrequently (e.g., 20-25% of the timed session), or 3) the word was derived from the approximation to a vowel (e.g., e for key, eh for red), consonant (e.g., k for king, b for bug), or novel diphthong (e.g., oy for boy) (Carbone, 2012). For Nolan, the word "Eat" was scored once during the assessment and selected for the study. This response appeared to be controlled by nonverbal stimulus (e.g., photo of a cake) and it was scored as a tact (by function). "Eat" did not occur under the presence of verbal stimulus (i.e., "Say, Eat" as an echoic) or when food was offered (EO for mand). For Ben, the word "Tissue" was derived from the consonant "t" and the word "shoe," and selected for the study. For Ken, who did not emit any words during his assessment, the word "Grandpa" was novel to the subject and was suggested by his parents. For Nelson, the word "Banana" was novel and was selected for this study. In the assessment, Nelson was able to reproduce the melody of the song "Old McDonald" by saying the consonant-vowel sound "Baba." This sound appeared to be controlled by a verbal stimulus (e.g., song) and it was denoted as an echoic. Therefore, the stimulus "banana" was derived from the consonant-vowel "baba." For Carla, the word "Hungry" was novel to the subject and also was suggested by the parent. For Eva, the word "Up" was derived from the low frequency sound, "Uh-oh" and was selected for this study. For Sam, the word "Eat" was novel and was suggested by his parent. For Tess, the word "Want" was novel and it was derived from the mand, "Wawa" (water) and suggested by her parent. Finally, for Matt, the word "Apple" was novel and selected for this study. Matt had the higher number of words of all the subjects but was unable to demonstrate echoics or tact “apple” during the assessment. The majority of the words were in a consonant-vowel-consonant-vowel form (e.g., mama, bike, baby).

Stimulus preference assessment. A paired-stimulus preference assessment (Fisher et al., 1992) was used to identify a hierarchy of four preferred food edibles, which were pre-selected based on parental consent and suggestions/recommendations, and the child’s dietary restrictions, if applicable. Edibles were presented in pairs, one pair at the time. The experimenter instructed the subject to pick one of the items. Subject was given 5 s to choose between the two edibles. If subject picked and consumed one of the edibles, a chosen response was scored. If a choice was not made the experimenter verbally prompted the subject to make a choice. If another 5 s has elapsed without a choice, the edibles were removed and both were scored as no chosen. Edibles were presented 3 times each over six trials. The number of times an edible was chosen out of the 3 presentations was multiplied by 100 to obtain a percentage of chosen. Table 2 displays the selection percentage for each edible included in the assessment for each subject. The highest selection percentage was denoted highly preferred (HP) and ranked in first place. The edible with moderate selection percentage was denoted moderately preferred (MP) and ranked in second place. The edible with the lowest selection percentage was denoted low preferred (LP) and ranked in the third place. For each subject, the HP edible was used during pairing and control conditions. These edibles included bunny crackers for Carla and Ken, Goldfish crackers for Nolan, Puff cereal for Ben and Tess, Cheerios cereal for Nelson, yogurt melts for Eva, Oreo cookies for Sam. The MP edibles were used in the adaptation session only. It is relevant to indicate that Matt only selected gummies (100%) and refused popcorn, Goldfish crackers and Puff cereals. However, his parent indicated that Goldfish was a second option; consequently, Goldfish crackers was ranked as MP edible. Similarly, Eva refused (e.g., verbally saying, “No, no, no") to engage in the preference assessment after 3 trials were conducted; consequently, Eva's choices were limited to yogurt melts. Eva selected peach flavored yogurt melts (67%), ranking as HP and vanilla flavored melts (33%), ranked as MP. Eva refused Puff cereals and Bunny crackers.

Settings and Materials

For Nolan, Carla, Eva, Sam, Tess, and Matt, sessions were conducted in the subjects' homes. Parents were invited to attend the training sessions but were not directly involved in any experimental procedures. For Nolan, Carla, and Eva, home sessions were conducted in the children's bedrooms. For Matt and Tess, sessions were conducted in the home's family rooms. For Sam, sessions were conducted in the experimenter's home. For Ken, Ben, and Nelson, sessions were conducted in a small college classroom. The experimenter provided toys (e.g., puzzle, ring stacker, shape form, trains and cars, and books) to all subjects across all sessions. These toys were intended for the entertainment of the children during training sessions and they were not assessed or otherwise identified to be preferred items. All sessions were recorded with a video camera 10-Sony HDR-CX675 High Definition Camcorder set on a fixed location of the room.

Apparatus. The apparatus consisted of a Homdox automatic food dispenser, a 60-cm by 60-cm folding table, a 10-cm by 15-cm black and white picture of a woman's face with an open mouth (use to visually prompt the subject's open mouth response), and an Insignia-Wave 2 portable Bluetooth speaker (see Figure 1). The auditory stimulus (i.e., the selected vocalization for the subject) was played from a computer to the speaker at 65 +5 dB (Vouloumanos et al., 2009). The experimenter pre-recorded the auditory stimulus (Petursdottir et al., 2011) by repeating the target word in an exaggerated and melodic fashion (i.e., motherese modeling) for the duration of a predetermined interval (e.g., 3, 6, or 12 s) to fill the CS-US gap. Using a pre-recorded sound facilitated consistent CS presentations across pairings and in relation to the ITI duration. At scheduled times, a small portion of edibles (e.g., Cheerios, Goldfish; based on the child's preference assessment results) was dispensed into a small tray of the dispenser for the child's consumption.

Response Measurement, Interobserver Agreement, and Procedural Integrity

The main dependent variables were target vocal response, approach-withdrawal ratio, and mouth opening. These responses were recorded via video for later data coding and analysis.

Target vocal response. This was defined as any sound or word produced by the subject that matched or shared similar acoustical features as the target auditory stimulus (e.g., selected sound or word). The rate of target vocal responses per minute was calculated by the sum of vocal responses produced by the subject divided by the total duration of the session (in minutes). Any vocal response separated by a 1-s silent interval from any other vocal response was scored as one response (Esch et al., 2009).

Approach-withdrawal response ratio. Within a 15-s onset of a pre-recorded target auditory stimulus (CS), an approach response was scored as one response when the subject oriented his or her head or walked toward the apparatus; otherwise, a withdrawal response was scored in the absence of any response described previously. The average approach-withdrawal response ratio was calculated by adding all approach responses per subjects and then dividing the sum by the total number of approach and withdrawal responses. A ratio near 1.00 indicated a strong approach toward the CS, and a ratio near 0.00 indicated strong withdrawal from the CS. A ratio near 0.50 indicated that the subject's response was not systematically controlled by the CS (Kaplan, 1984).

Mouth opening. The opening of the subject's mouth was scored as one response if there was a gap or separation of the subject's mouth prior to any vocalization or target vocal response. The rate of mouth openings per minute was calculated by the sum of mouth openings divided by the total duration of the session (in minutes).

Two independent observers collected data on the dependent variables after sessions from videos. The observers did not collect data during sessions. Interobserver agreement (IOA) was assessed for 50% of sessions for all subjects by using exact count-per-interval IOA (Cooper, Heron, & Heward, 1987), and was calculated by dividing the number of intervals of 100% IOA within each 1-min interval by the total number of intervals for each session and multiplying by 100. Agreements were defined as both observers indicating the same number of occurrences or nonoccurrences of behavior in each interval, whereas a disagreement occurred when only one of the two observers recorded the behavior in a corresponding interval. Mean IOA was 93% (range, 86% to 100%) for Ave, 98% (range, 95% to 100%) for Nolan, 94% (range, 88% to 100%) for Carla, 98% (range, 95% to 100%) for Ben and Sam, and 100% for Ken, Nelson, Tess, and Matt.

Design and Procedure

A pairwise design was used to demonstrate experimental control (Iwata, Duncan, Zarcone, Lerman, & Shore, 1994). Pairing (test) and control conditions were sequentially ordered and alternated across sessions. Three subjects were randomly assigned to one of three experimental groups. A test and a control condition were conducted once per week within one single appointment, but there were occasions where only one condition was conducted due to time constraint, child's illness or technical issues (e.g., the video camera stopped working). The experimenter then arranged the missed condition in a make-up session. Table 3 displays the distribution of subjects per groups, target vocalizations, and type of vocalization selected.

Experimental sessions were arranged close to or before any scheduled snacks or meals, if possible. A session was suspended by 10 min if a subject acted tired or began to cry. If the child continued to be irritable beyond the 10-min break, the session was terminated and a make-up session was scheduled. During the experiment, the parents were invited to be present in the same room but were instructed to avoid any verbal interaction with their child. Parents were encouraged to ask for the assistance of the experimenter when needed or to stop the session at their discretion. Subject-experimenter interactions were limited to only actions for preserving the child's safety in all and across conditions to control influence of familiar/unfamiliar pairing agent (Petursdottir & Lepper, 2015).

Experimental conditions.

***Adaptation*.** Each subject was trained to pick up the edibles from the apparatus tray on the first session only. The experimenter placed 3 pieces of a MP edible (according to preference assessment results) in the apparatus tray. The experimenter waited 15 s for the subject to consume the edibles. If there was no response, a verbal (e.g., "There is food here") and gestural prompt (e.g., pointing to the tray) were provided to initiate the response. After the subject consumed the food, the experimenter set a timer and dispensed edibles into the tray after an average of 30 s (range: 10- 40 s; Esch et al. 2009). The subject was able to walk around the room and play with toys without restriction. The adaptation phase ended when the subject was able to pick up the food from the tray within 5 s of food presentation across five successful trials. Following this phase, the subject received a 5-min break and the experimenter arranged the apparatus for the following condition.

***Pairing*.** The experimenter began a pairing session by presenting a pre-recorded target auditory stimulus (CS-onset) selected for the child for the entire duration of such stimulus (ISI).

On the ISI termination, the experimenter then pressed a button on the apparatus to dispense HP edibles (US) into the tray for the child's consumption. If the subject did not consume the food within 5 s of US presentation (as per recommendation in Esch et al., 2009), the experimenter scooped the food and replaced it into the apparatus. Then, the experimenter set a digital timer for predetermined ITI duration before presenting the CS and beginning the next trial. Tables 4 through 6 display the values of ITIs and ISIs across trials and groups. Across groups, the average ITI duration was varied, with values of 15 s (range: 15-25 s), 30 s (range: 20-40 s), and 60 s (range: 40-80 s). Inter-stimulus-interval (ISI) duration was also varied to 1:5 ratio of the ITI value. For example, in a trial with a15-s ITI, the ISI value was equal to 3 s. Each subject received 60 pairing trials. In the 15-s ITI group and the 30-s ITI group, 30 pairing trials were conducted per session for a total of two pairing sessions. In the 60-s ITI group, pairing sessions included 20 trials for a total of three pairing sessions. Session lengths varied but never exceeded 35 min.

Control. The experimenter began a control session by setting two independent digital timers (VT 15 s, 30 s, and 60 s) that cued the presentation of CS and US randomly and independently (i.e., truly random control; Rescorla, 1967), which led to a 50/50 distribution of stimulus presentations. As suggested this control procedure seemed more adequate than a delay control procedure commonly used in SSP research because it is based on the idea that the occurrence of the CS gives no information about the occurrence of the US; consequently, it eliminates the new contingency established by the delay control procedure (Rescorla,1976).

Each subject received 60 stimulus presentations. In the 15-s and 30-s groups, 30 stimulus presentations were presented per session for a total of two control sessions. In the 60-s group, 20 stimulus presentations were presented per session for a total of three control sessions. The experimenter arranged the stimulus presentations independently from the occurrence of CR. If the subject did not consume the food within 5 s of US presentation, the experimenter scooped the food and replaced it into the apparatus. If at any point the presentation of CS and US coincided, the experimenter scored the trial as a pairing.

Table 7 displays subjects' session durations across pairing and control conditions. In the 15-s ITI group, average session duration was 22 min for pairing and 21 min for control. In the 30-s ITI group, average session durations were 26 min for pairing and 25 min for control. In the 60-s ITI group, average session durations were 31 min for pairing and 27 min for control.

Results

Figures 2 through 4 display results for each subject, indicating that higher rates of vocal responses were observed in the pairing conditions. Figure 2 displays the results of each of the subjects of the 15-s ITI group. For Carla (top panel), mean increases from control to SSP were 0.10. Carla's vocal responses increased initially in both control and pairing and subsequently decreased to 0 in both conditions. However, the rate of vocal responses was higher in the pairing condition (0.25 rpm) than in the control condition (0.05 rpm). For Ken (middle panel), mean increases from control to pairing were 1.46. Ken's vocal responses increased significantly in the first pairing session (2.65 rpm) and then decreased in the second pairing session (0.32 rpm). For Matt (bottom panel), mean increases from control to pairing were 0.46 rpm. Matt's vocal responses increased in both pairing and control conditions, but rate of responses was significantly higher in SSP (0.7 rpm) than control (0.24 rpm).

Figure 3 displays the results of each of the subjects of the 30-s ITI group. For Eva (top panel), mean increases from control to pairing were 0.08 rpm. Eva's vocal responses significantly increased across pairing sessions (0.05 rpm in Session 2, 0.11 rpm in Session 4) but none of the vocal responses increased in the control sessions. For Nolan (middle panel), mean increases from control to pairing were 0.08 rpm. Nolan's vocal responses increased in the first pairing session to 0.3 rpm but subsequently decreased to 0.05 rpm in the second pairing session. As in the pairing condition, vocal responses also increased by 0.2 rpm in the control sessions. Interestingly, as the rate of vocal responses decreased in the pairing sessions, rates increased in the control sessions. For Tess (bottom panel), mean increases from control to pairing were 0.25 rpm. Tess's vocal responses slightly increased in both pairing sessions (0.2 rpm in session 1 and 0.3 rpm in session 3), and none of the vocal responses occurred in the control sessions.

Figure 4 shows the results of each subject in the 60-s ITI group. For Ben (top panel), mean increases from control to pairing were 0.18 rpm. Ben's vocal responses increased initially and later plunged to 0 across all sessions. For Nelson (middle panel), mean increases from control to pairing were 0.34 rpm. Nelson's vocal responses increased in rate progressively over pairing sessions (0.13, 0.4, and 0.5, respectively), and none occurred in the control sessions. For Sam (bottom panel), mean increases from control to pairing were 0.26 rpm. Sam's vocal responses increased initially (0.5 rpm) and then decreased and remained at the same level across pairing sessions 4 and 6 (0.2 and 0.21, respectively). Vocal responses remained at low levels across control sessions (0.06, 0, and 0.06, respectively).

Figure 5 displays mean vocal responses per groups during pairing and control sessions. Mean increases from control to pairing were 0.67 in the 15-s ITI group; 0.14 in 30-s ITI group; and 0.26 in the 60-s ITI group, which indicates 15-s ITI group showed greater increases of vocal responses in comparison with other groups. Mean vocal responses in this group were 0.77 rpm (SEM = 0.394) in the pairing condition and 0.1 rpm (SEM = 0.063) in the control. Responses were significantly higher in the first session than in the other conditions but decreased in the subsequent session to levels similar to the rates in the 30-s ITI and 60-s ITI. Subjects in the pairing 30-s ITI group averaged 0.17 rpm (SEM = 0.04) and 0.03 rpm (SEM = 0.033) in the control. Subjects in the pairing 60-s ITI group averaged 0.27 rpm (SEM = 0.04) and 0.01 rpm (SEM = 0.013) in the control condition.

Figure 6 depicts mean vocal responses and mean approach-withdrawal ratios per groups across sessions. As indicated above, subjects in the 15-s ITI condition showed higher rates of vocalizations than subjects in 30-s ITI and 60-s ITI conditions but vocal responses displayed a downward trend as withdrawal responses became closer to 0. Subjects in the 30-s ITI and 60-s ITI demonstrated moderate and stable rates of vocal responses and exhibited approach responses above 0.6; however, only subjects in the 60-s ITI showed approach responses above 0.8.

Figure 7 depicts a within session analysis of mean approach-withdrawal response ratios across 10-trial block sessions. Mean approach-withdrawal responses were 0.45 (SEM = 0.0577) in the 15-s ITI condition, 0.62 (SEM = 0.112) in the 30-s ITI condition, and 0.80 (SEM = 0.070) in the 60-s ITI condition. Latter group exhibited approach responses to the CS above 0.7 more consistently than other groups; however, all groups showed values above 0.70 in the first two sessions. Subjects in the 15-s ITI withdrew from the CS more often than subjects in the other groups as values plunged below 0.5 after the second session.

Mouth opening was a secondary dependent variable measured in this study. However, none of the subjects displayed mouth opening (as defined above) in any of the conditions. Therefore, results for rate of mouth opening are not shown or discussed at this time.

Discussion

The results of this study support and extend the findings of previous research in SSP by showing that young children with various levels of language skills can acquire novel vocal responses through SSP. All subjects showed higher rates of vocalizations in the pairing conditions, with higher rates of vocalizations in the 15-s ITI group. The implications of these results seem substantial because subjects were able to learn new words without the use of contingent reinforcement, echoic training or prompting (Sundberg et al., 1996). Particularly Nelson, who has never spoken any words before the training, was able to produce the target response on the first pairing session and responding levels continued to increase in subsequent pairing sessions, but never in the control condition. These findings add to the relevance of SSP as a clinical procedure that can aid nonvocal children who would not benefit from vocal shaping or echoic training due to their inability to produce or imitate initial sounds. Moreover, previous research (Carroll & Klatt, 2008; Miguel et al., 2002) speculated that the type of vocalizations (i.e., novel versus existing vocalizations) might have differential effects on SSP. In fact, Carroll and Klatt (2008) reported it might be easier to increase existing rather than novel vocalizations. However, current study showed eight of nine subjects learned novel vocalizations as well as the one subject whose existing low frequency vocalization was selected.

The current study tested the effects of varying the duration of the ITI and the ISI in the establishment of human vocalizations based on recommendations from responding research (e.g., autoshaping; Brown & Jenkins, 1968; Kaplan, 1984) to resolve some procedural limitations of SSP (da Silva and Williams, under review).We hypothesized the relative duration of ITI and ISI might have a critical impact on conditioning of human vocalizations as has been observed in the conditioning of other species' responses (e.g., Jenkins & Moore, 1973; Sidman & Fletcher, 1968). To date, no SSP studies have evaluated the isolated effect of these variables, but rather, had focused on other factors (e.g., participants' age and diagnosis, type of pairing, number of sound-presentations, type of reinforcer, and control procedures). In agreement with the findings of Kaplan's (1984) study human and nonhuman subjects were similarly sensitive to the relative values of ISI and ITI, and subjects showed a strong approach response in the longer ITI condition. As we predicted, subjects in the 60-s ITI condition exhibited a strong approach response to the CS. In fact, longer ITI duration might be preferable due to increasing the chances for subjects to acquire a strong sign tracking behavior (i.e., orientating to the CS). However, subjects in the 30-s ITI group demonstrated mild approach response to the CS and did not perform as expected, which indicates that this condition can also be appropriate if shorter pairing trials are desired. Additionally, subjects in the 15-s ITI condition failed to acquire any clear tendency to approach or withdraw from the CS, and these subjects performed similarly to the nonhuman subjects in the Kaplan's 60-s ITI condition. These findings are relevant to explain SSP failure to condition vocalizations for some or all participants in previous SSP studies (Carroll & Klatt, 2008; Esch et al., 2005; Miguel et al., 2002; Normand & Knoll, 2006; Stock et al., 2008) in which ITI duration was shorter than 30 s (e.g., 0-20 s). This information should be carefully considered in future SSP research. Nevertheless, Ken (a subject in first group) showed the highest rates of vocalizations per min in the pairing conditions of all subjects among groups. This difference could be due to the subject's individual history of reinforcement. Another possible contributing factor was the use of motherese modeling to enhance the CS by producing such with a melodic voice. Shillingsburg et al. (2015) had reported motherese modeling can be effective to increase rates of vocal responses.

These findings provide further empirical support for varying the time between trials. Previous research (Barry et al., 2019, Esch et al., 2009, Miliotis et al., 2012, and Rader et al., 2014) had demonstrated success using 5-30 s variable ITI durations. Varying the time between trials seems important because eliminates temporal predictability of the presentation of preferred items (i.e., US is unexpected).

Another aspect evaluated in this study was the use of a truly random control (Rescorla, 1967). Many past studies had assessed the utility of an omission control procedure (i.e., correction delay) to control for adventitious reinforcement of responses but results had been mixed. In this study, seven of nine subjects showed differentiated low to no levels of responding in the control conditions in comparison to the pairing condition, but all acquired the target vocal responses. However, Matt and Nolan exhibited moderate rates of vocalizations in the control conditions. First, we speculate these subjects might have failed to discriminate changes in the conditions; consequently, future research should include stimulus control procedures (e.g., colored cards) to signal to the subject the shift to the next condition. Second, Matt and Nolan might have exhibited target vocal responses because the presence of the EO for food became strong when the delivery of the US no longer followed the CS, and such EO could have evoked the response as a mand (as we observed the children waiting in front of the apparatus to receive the food while they emitted the words). Future research should apply this procedure with other forms of reinforcement (e.g., physical touch, smiles, eye contact) to further evaluate its utility in SPP.

An additional level of control considered was using an apparatus to present stimuli independent from a human trainer. We hypothesized the apparatus might have helped to standardize the delivery of the US and control for other sources of reinforcement (e.g., smiles, head nods, eye contact, etc.) mediated by a trainer. Therefore, we made deliberated attempts to guarantee that the CS and no other stimulus was paired with the US.

Another interesting aspect of this study is how some of the subjects emitted newly paired vocal responses sporadically as a form of vocal play at different times throughout the day (especially for Ben). This finding appears to be consistent with results reported in previous studies (e.g., Smith et al., 1996; Sundberg et al., 1996). For example, the repeated pairing of the word with a preferred edible possibly made Ben's vocalizations more sensitive to automatic reinforcement when the child heard himself produce sounds or words that share similar acoustic features with the paired word (Petursdottir & Mellor, 2017; Shillingsburg et al., 2015; Smith et al., 1996).

The current study included several limitations. First, the use of only one video camera limited the quality of the video and reduced the number of opportunities to observe subject's engagement in opening mouth as the subjects moved frequently around the experimental area during training sessions. Further research should consider including both video and in-vivo recording of target responses. Second, motherese modeling was included as a procedure to enhance CS salience but its isolated effects were not measured nor control. Third, the duration of training sessions did not exceed 35 minutes, but some subjects acted tired most often as the training session went along. It might be best practice to conduct shorter sessions with fewer trials spaced out across days. It is still a relevant question to determine the optimal number of trials to observe an effect of SSP. For example, Myers (1981) included a criterion of 15 consecutive trials with CR before discontinuing pairing procedures. Establishing a learning criterion can also guide practitioners' decision to begin other procedures (e.g., mand or echoic training) to capture emerging vocalizations. Future research should evaluate ways to increase the efficiency of SSP procedure in clinical settings.

The current study attests that SSP effectiveness relies, at least in part, on the relative temporal contiguity of events. Varying the duration of the ISI to respect of the ITI produced positive outcomes among subjects across groups but showed that longer ITIs (e.g., 30-60 s) are preferable to sustain levels of responding.

References

Barry, L., Holloway, J., & Gunning, C. (2019). An investigation of the effects of a parent delivered stimulus-stimulus pairing intervention on vocalizations of two children with Autism Spectrum Disorder. *The Analysis of Verbal Behavior*, *35*, 57-73. https://doi.org/10.1007/s40616-018

Bloom, K., & Esposito, A. (1975). Social conditioning and its proper control procedures. *Journal of Experimental Child Psychology, 19,* 209-222. http://dx.doi.org/10.1016/0022-0965(75)90085-5

Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior, 11,* 1-8. http://doi.org/10.1901/jeab.1968.11-1

Carbone, V. (2012). Increasing speech sound production of children with autism. Presentation. Massachusetts.

Carroll, R. A., & Klatt, K. P. (2008). Using stimulus-stimulus pairing and direct reinforcement to teach verbal behavior to young children with autism. *The Analysis of Verbal Behavior, 24,* 135-146. https://doi.org/10.1007/BF03393062

Cherry, K. (2018, September 16). *Communication Milestones in Child Development*. Retrieved fromhttps://www.verywellmind.com/communication-milestones-2795110

Children's Hospital of Orange County (2018). *Child development: Ages and stages* CHOC.org. Retrieved from https://www.choc.org/neuroscience/developmental-services/ages-stages/#2-years

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. https://doi.org/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. https://doi.org/10.1007/BF03393009

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Ownes, 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. https://doi.org/10.1901/jaba.199225-491

Gamzu, E., & Schwam, E. (1974). Autoshaping and automaintenance of a key-press response in squirrel monkeys. *Journal of the experimental analysis of behavior*, *21*, 361–371. doi:10.1901/jeab.1974.21-361

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between bird song and speech. *Proceedings of the National Academy of Sciences of the United States of America, 100,* 8030-8035. https://doi.org/10.1073/pnas.1332441100

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

Iwata, B. A., Duncan, B. A., Zarcone, J. R., Lerman, D. C., & Shore, B. A. (1994). A sequential test-control methodology for conducting functional analyses of self-Injurious behavior. *Behavior Modification*, *18*, 289-306. https://doi.org/10.1177/01454455940183003

Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior, 20*, 163-181. https://doi.org/10.1901/jeab.1973.20-163

Kaplan, P. S., & Hearst, E. (1982). Bridging temporal gaps between CS and US in autoshaping: Insertion of other stimuli before, during, and after CS. *Journal of Experimental Psychology, 8*, 187-203. http://dx.doi.org/10.1037/0097-7403.8.2.187

Kaplan, P. S. (1984). Importance of relative temporal parameters in trace autoshaping: From excitation to inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 113-126. http://dx.doi.org/10.1037/0097-7403.10.2.113

Koegel, R. L., Koegel, L. K., & Surratt, A. (1992). Language intervention and disruptive behavior in preschool children with autism. *Journal of Autism and Developmental Disorders*, *22*, 141-153. https://doi.org/10.1007/BF01058147

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. https://doi.org/10.1002/jaba.55

Locurto, C., Terrace, H. S., & Gibbon, J. (1976). Autoshaping, random control, and omission training in the rat. *Journal of the Experimental Analysis of Behavior, 26,* 451-462. https://doi.org/10.1901/jeab.1976.26-451

Lucas, G. A., & Wasserman, E. A. (1982). US duration and local trial spacing affect autoshaped responding. *Animal Learning & Behavior, 10,* 490-498. http://dx.doi.org/10.3758/BF03212289

McLaughlin, S. F. (2010). Verbal behavior by B.F.Skinner: Contributions to analyzing early language learning. *The Journal of Speech-Language Pathology and Applied Behavior Analysis, 5,* 114-131. http://dx.doi.org/10.1037/h0100272

Miguel, C. F., Carr, J. E., & Michael, 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. https://doi.org/10.1007/BF03392967

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. https://doi.org/10.9101/jaba.2012.45-809

Myers, A. M. (1981). *Autoshaping Infant Vocalizations.* Dissertation, Utah State University. Psychology Commons.

National Autism Center. (2018). *Autism Fact Sheet.* Retrieved from https://nationalautismassociation.org/resources/autism-fact-sheet/

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. https://doi.org/10.1007/BF03393028

Palmer, D.C. (2018). The role of automatic reinforcement in shaping speech*.* Presentation. Pennsylvania.

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. https://doi.org/10.1007/BF03393091

Petursdottir, A. I., & Lepper, T. I. (2015). Inducing novel vocalizations by conditioning speech sounds as reinforcers. *Behavior Analysis in Practice, 8,* 223-232. https://doi.org/10.1007/s4617-015-0088-6

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

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

Rescorla, R. A. (1984). Associations between Pavlovian CSs and context. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 195-204. http://dx.doi.org/10.1037/0097-7403.10.2.195

Ricci, J. (1973). Key pecking under response-independent food presentation after long simple and compound stimuli. *Journal of the Experimental Analysis of Behavior, 19,* 509-516. https://doi.or/ 10.1901/jeab.1973.19-509

Rosales- Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis, 30,* 533-544 https://doi.org/10.1901/jaba.1997.30-533

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. https://doi.org/10.1007/s40616-015-0042-2

Sidman, M., & Fletcher, F. G. (1968). A demonstration of auto-shaping with monkeys. *Journal of the Experimental Analysis, 11*, 307-309. http://dx.doi.org/10.1901/jeab.1968.11-307

da Silva, S. P., & Williams, A. M. (2019). Analogies in stimulus pairing: Procedural considerations based on autoshaping. Manuscript submitted for publication.

Skinner, B. F. (1957). *Verbal behavior.* New York: Appleton-Century-Crofts.

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. http://dx.doi.org/10.1007/BF03392905

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. http://dx.doi.org/10.1007/BF03393061

Sundberg, M. L. (2008). VB-MAPP: *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. https://doi.org/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.

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities.* Pleasant Hill, CA: Behavior Analysts, Inc.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*, 1245-1260. https://doi.org/10.1044/jslhr.4006.1245

Vaughan, M. E., & Michael, J. L. (1990). Automatic reinforcement: An important but ignored concept. *Behaviorism, 10*, 217-227

Vouloumanos, A., Druhen, M. J., Hauser, M. D. & Huizink, A. T. (2009). Five-month-old infants' identification of the sources of vocalizations. *Psychological and Cognitive Sciences, 106,* 18867-18872. https://doi.org/10.1073/pnas.0906049106

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. https://doi.org/10.1007/BF03393050

Wilcove, W. G., & Miller, J. C. (1974). CS-UCS presentations and a lever: Human autoshaping. *Journal of Experimental Psychology, 103,* 868-877. https://doi.org/10.1037/h0037388

Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: sustain pecking despite contingent non-reinforcement. *Journal of Experimental Behavior Analysis, 12,* 511-520.

https//:doi: 10.1901/jeab.1969.12-511

Wu, Z., & Gros-Louis, J. (2016). Infant's prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language, 34,* 72-90. https://doi.org/10.1177/0142723714521925

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.

Table 1

VB-MAPP Results Across Subjects

Verbal Operant	Matt	Tess	Sam	Nelson	Ben	Nolan	Ave	Ken	Carla
Mand	4	3	1	0	0	3	2	1	4
Tact	2	4	1	0	0	3	2	0	5
LR	5	5	5	5	5	5	5	5	5
Echoic	2	5	1	0	4	2.5	4	1	4
Total	13	17	8	5	9	13.5	13	7	18

Note. Individual skill scores were obtained by scoring the subject response based on the criteria identified in each section of the specific milestone scoring form. A response was scored based on three options: 0, ½, or 1. Then, total scores were obtained by adding up all the points acquired by the subject for each skill area.

Table 2

Subjects' Stimulus Preference Assessment

Subject	Highest Preferred Stimulus	Moderately Preferred Stimulus	Least Preferred Stimulus	No selection
Carla	Bunny crackers (50%)	Goldfish crackers (33%)	Puff cereal (17%)	Cheerios (0%_
Ben	Puff cereal (50%)	Yogurt melts (33%)	Bunny crackers (17%)	Goldfish crackers (0%)
Eva	Peach yogurt melts (67%)	Vanilla yogurt melts (33%)		Puff cereal and bunny crackers (0%)
Ken	Bunny crackers (50%)	Goldfish crackers (33%)	Cheerios (17%)	Puff cereal (0%)
Sam	Oreo cookies (50%)	Puff cereal (33%)	Bunny cracker (17%)	Goldfish crackers (0%)
Matt	Fruit gummies (100%)	*Goldfish crackers		Popcorn and Puff cereal (0%)
Tess	Puff cereal (50%)	Bunny crackers (33%)	Cheerios (17%)	Goldfish crackers (0%)
Nolan	Cheerios (50%)	Goldfish, bunny, and cheese crackers (17% respectively)		
Nelson	Cheerios (50%)	Goldfish crackers (33%)	Puff cereal (17%)	Yogurt melts (0%)

Note. Selection percentages rounded to the nearest whole number are shown below the stimulus within the parentheses. Highest preferred stimulus was used as US during pairing and control conditions. Moderately preferred stimulus was used as preferred item during adaption. For Matt, *Goldfish crackers were arbitrary selected as MP based on parental suggestion.

Table 3

Distribution of Subjects per Group and Selected Vocal Responses

Group	Subject	Type of Vocal	Target Vocal
15-s ITI	Carla	Novel	Hungry
	Ken	Novel	Grandpa
	Matt	Novel	Apple
30-s ITI	Eva	Novel	Up
	Nolan	Low frequency	Eat
	Tess	Novel	Want
60-s ITI	Ben	Novel	Tissue
	Nelson	Novel	Banana
	Sam	Novel	Eat

Table 4

Distribution of ITI and ISI Values Across Pairing Trials in 15-s ITI group

Trial	ISI	ITI	Trial	ISI	ITI
1	3	15	31	5	25
2	5	25	32	4	20
3	4	20	33	3	15
4	4	20	34	5	20
5	5	25	35	4	20
6	5	25	36	3	15
7	5	25	37	5	25
8	4	20	38	5	25
9	4	20	39	5	25
10	5	25	40	4	20
11	3	15	41	3	15
12	5	25	42	3	15
13	4	20	43	5	25
14	5	25	44	4	20
15	5	25	45	3	15
16	3	15	46	5	25
17	4	20	47	5	25
18	5	25	48	4	20
19	3	15	49	5	25
20	3	15	50	3	15
21	4	20	51	5	25
22	5	25	52	4	20
23	5	25	53	4	20
24	5	25	54	5	25
25	3	15	55	5	25
26	4	20	56	5	25
27	5	25	57	4	20
28	3	15	58	4	20
29	4	20	59	5	25
30	5	25	60	3	15

Note. Average ITI duration was varied with values of 15-s ITI (range: 15-25 s) and the average ISI duration was varied with values 3-s ISI (range: 3-5 s)

Table 5

Distribution of ITI and ISI Values Across Pairing Trials in 30-s ITI group

Trial	IS1	ITI	Trial	S1	ITI
1	6	30	31	6	30
2	5	25	32	5	25
3	6	30	33	6	30
4	7	35	34	7	35
5	7	35	35	6	30
6	4	20	36	7	35
7	6	30	37	4	20
8	4	20	38	5	25
9	7	35	39	5	25
10	8	40	40	7	35
11	4	20	41	7	35
12	7	35	42	5	25
13	6	30	43	7	35
14	4	20	44	5	25
15	5	25	45	6	30
16	6	30	46	5	25
17	5	25	47	4	20
18	7	35	48	6	30
19	5	25	49	7	35
20	7	35	50	4	20
21	7	35	51	8	40
22	5	25	52	7	35
23	5	25	53	4	20
24	4	20	54	6	30
25	7	35	55	4	20
26	6	30	56	7	35
27	7	35	57	7	35
28	6	30	58	6	30
29	5	25	59	5	25
30	6	30	60	6	30

Note. Average ITI duration was varied with values of 30-s ITI (range: 20-40 s) and the average ISI duration was varied with values 6-s ISI (range: 4-8 s)

Table 6

Distribution of ITI and ISI Values Across Pairing Trials in 60-s ITI group

Trial	IS1	ITI	Trial	IS1	ITI	Trial	IS1	ITI
1	10	50	21	16	80	41	10	50
2	13	65	22	15	75	42	13	65
3	15	75	23	13	65	43	15	75
4	10	50	24	10	50	44	10	50
5	11	55	25	12	60	45	11	55
6	13	65	26	13	65	46	13	65
7	10	50	27	11	55	47	10	50
8	16	80	28	10	50	48	16	80
9	11	55	29	11	55	49	11	55
10	10	50	30	16	80	50	10	50
11	16	80	31	10	50	51	16	80
12	11	55	32	11	55	52	11	55
13	10	50	33	16	80	53	10	50
14	11	55	34	10	50	54	11	55
15	13	65	35	13	65	55	13	65
16	12	60	36	11	55	56	12	60
17	10	50	37	10	50	57	10	50
18	13	65	38	15	75	58	13	65
19	15	75	39	13	65	59	15	75
20	16	80	40	10	50	60	16	80

Note. Average ITI duration was varied with values of 60-s ITI (range: 40-80 s) and the average ISI duration was varied with values 12-s ISI (range: 8-16 s)

Table 7

Subjects' Session Durations Across Pairing and Control Conditions

Condition	Subject	Pairing	Control
15-s ITI	Carla	Session 2 (22 min)	Session 1 (22 min)
		Session 4 (22 min)	Session 3 (21 min)
	Ken	Session 1 (21 min)	Session 2 (21 min)
		Session 3 (22 min)	Session 4 (22 min)
	Matt	Session 2 (21 min)	Session 1 (21 min)
		Session 4 (20 min)	Session 3 (21 min)
30-s ITI	Eva	Session 2 (26 min)	Session 1 (25 min)
		Session 4 (25 min)	Session 3 (25 min)
	Nolan	Session 2 (28 min)	Session 1 (25 min)
		Session 4 (27 min)	Session 3 (25 min)
	Tess	Session 1 (25 min)	Session 2 (25 min)
		Session 3 (27 min)	Session 4 (25 min)
60-s ITI	Ben	Session 1 (31 min)	Session 2 (28 min)
		Session 3 (30 min)	Session 4 (28 min)
		Session 5 (30 min)	Session 6 (27 min)
	Nelson	Session 1 (32 min)	Session 2 (27 min)
		Session 3 (30 min)	Session 4 (26 min)
		Session 5 (30 min)	Session 6 (26 min)
	Sam	Session 1 (32 min)	Session 2 (27 min)
		Session 3 (30 min)	Session 4 (28 min)
		Session 5 (32 min)	Session 6 (26 min)

Figure 1. Apparatus used in pairing and control conditions.

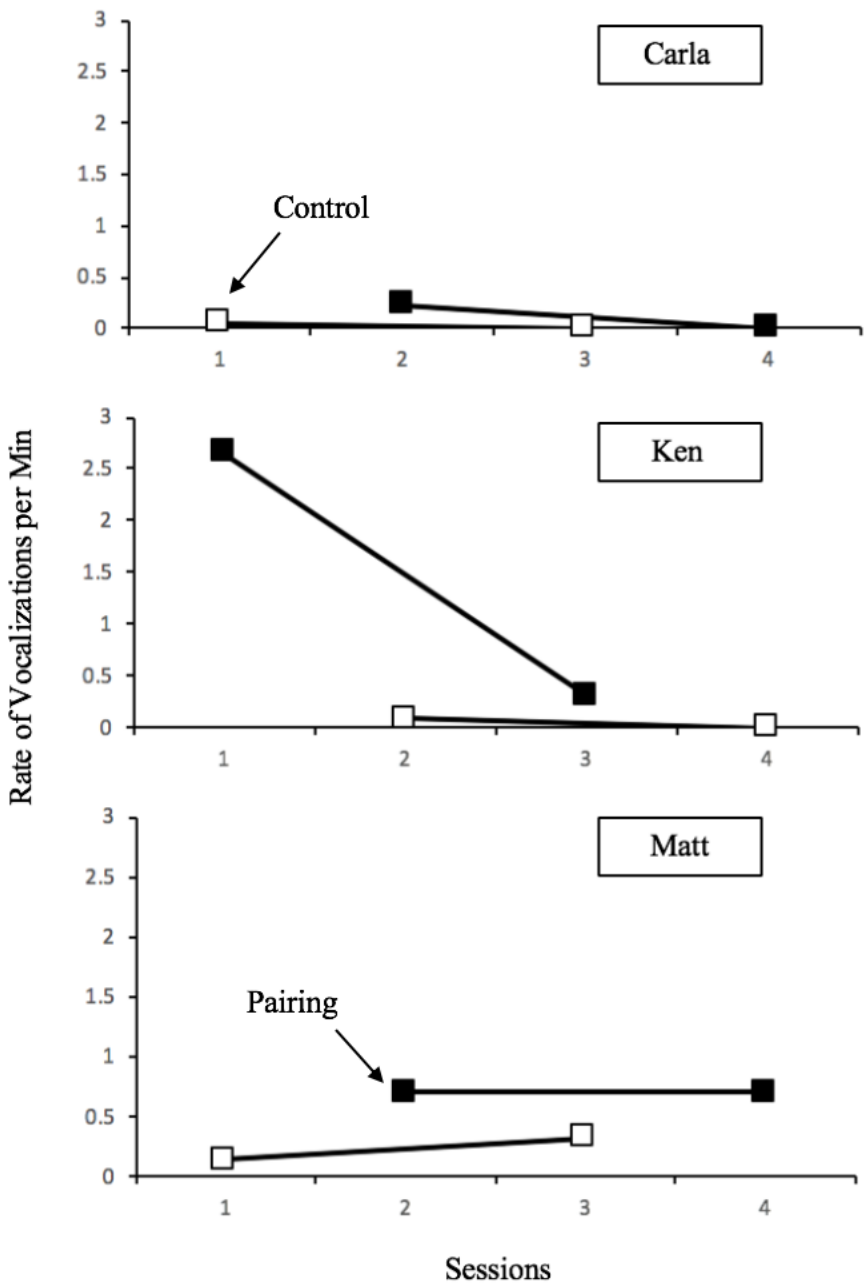

Figure 2. Rate of target vocalizations for Carla, Ken and Matt across pairing (closed squares) and control (open squares) conditions in 15-s ITI group.

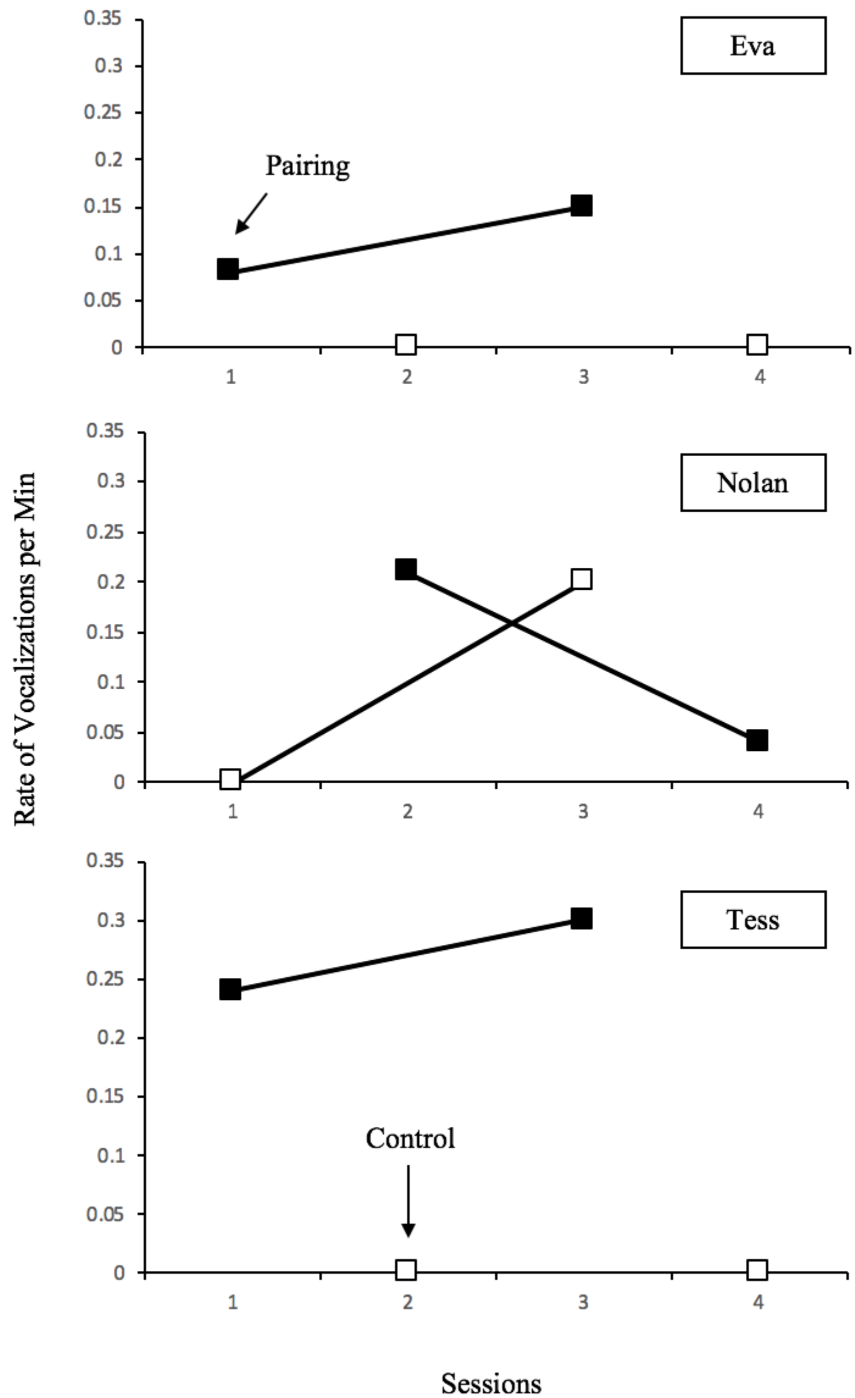

Figure 3. Rate of target vocalizations for Eva, Nolan and Tess across pairing (closed squares) and control (open squares) conditions in 30-s ITI group.

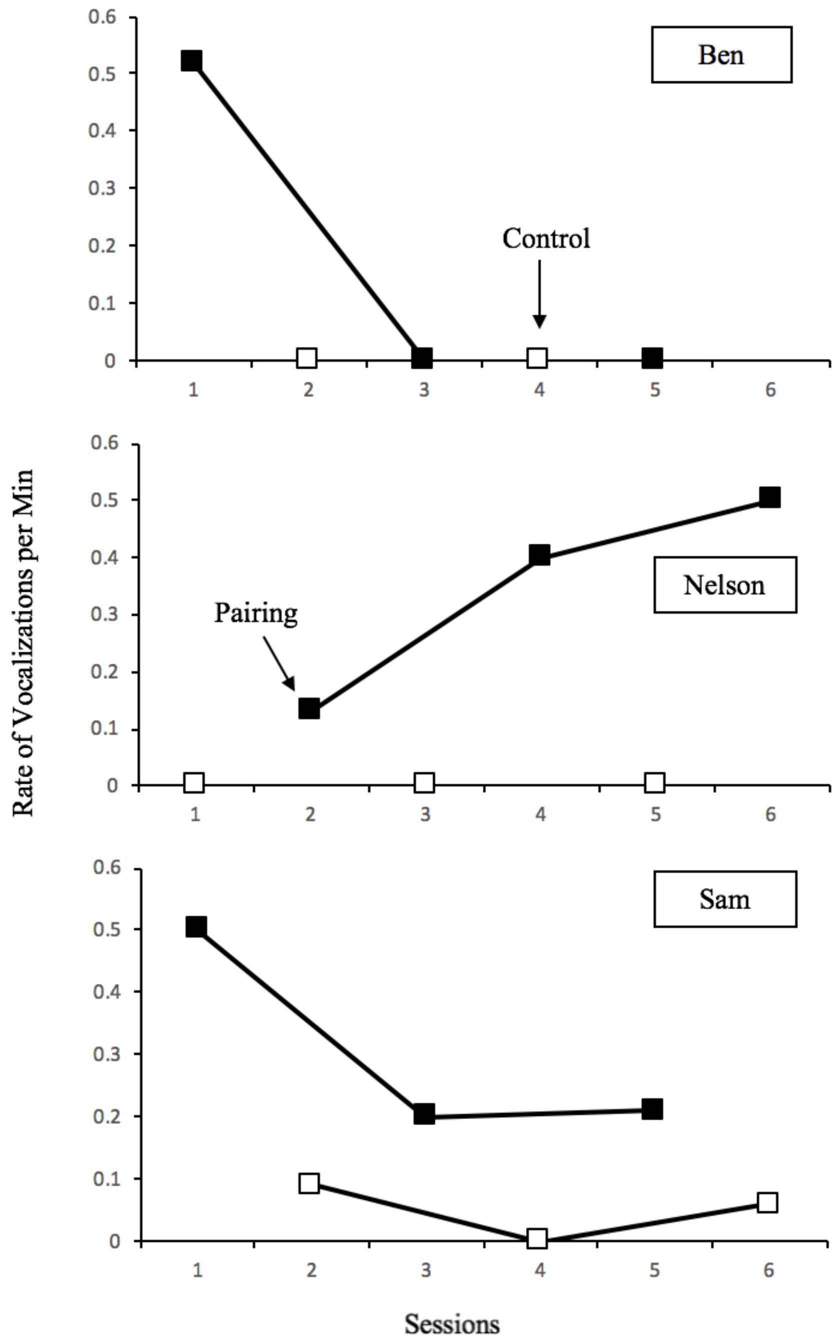

Figure 4. Rate of target vocalizations for Ben, Nelson and Sam across pairing (closed squares) and control (open squares) conditions in 60-s ITI group.

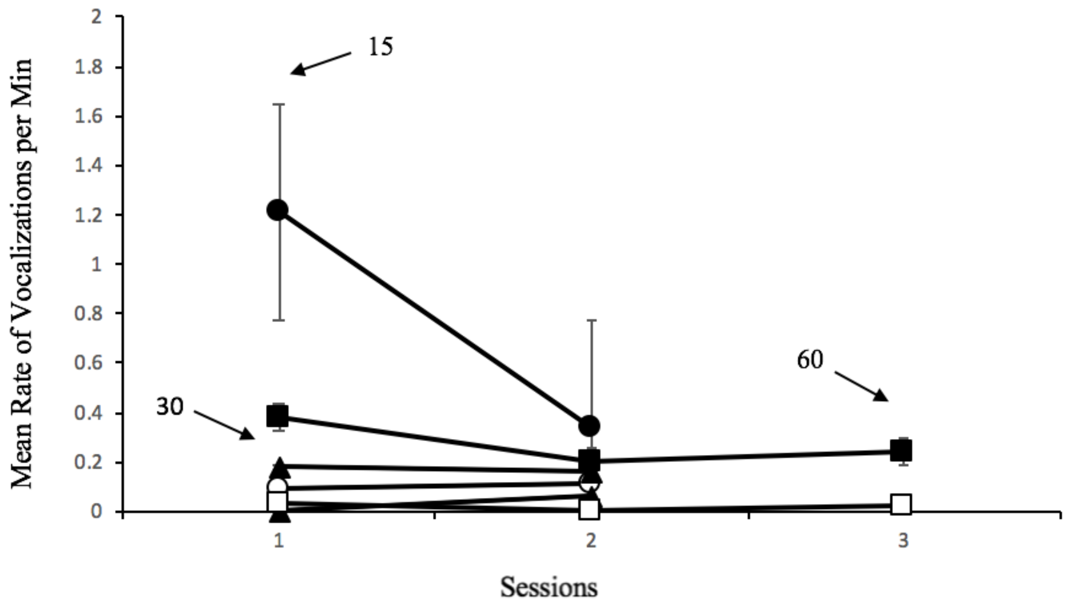

Figure 5. Mean rate of vocal responses across pairing sessions for subjects in the 15-s ITI (closed circles), 30-s ITI (closed triangles), and 60-s ITI (closed squares) conditions. ***5.*** Mean rate of vocal responses across control sessions for subjects in the 15-s ITI (open circles), 30-s ITI (open triangles), and 60-s ITI (open squares) conditions Error bars represent standard errors.

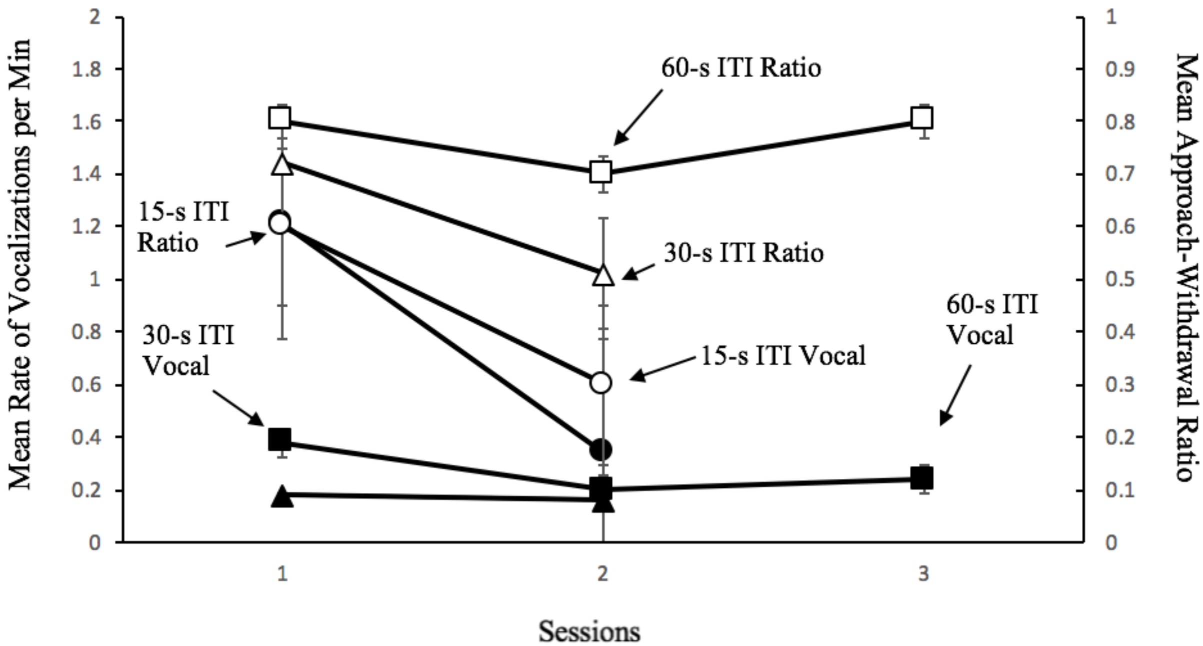

Figure 6. Mean rate of vocal responses across pairing sessions for subjects in the 15-s ITI (closed circles), 30-s ITI (closed triangles), and 60-s ITI (closed squares) conditions. Mean approach-withdrawal ratios across pairing sessions for subjects in the 15-s ITI (open circles), 30-s ITI (open triangles), and 60-s ITI (open squares) conditions. Error bars represent standard errors.

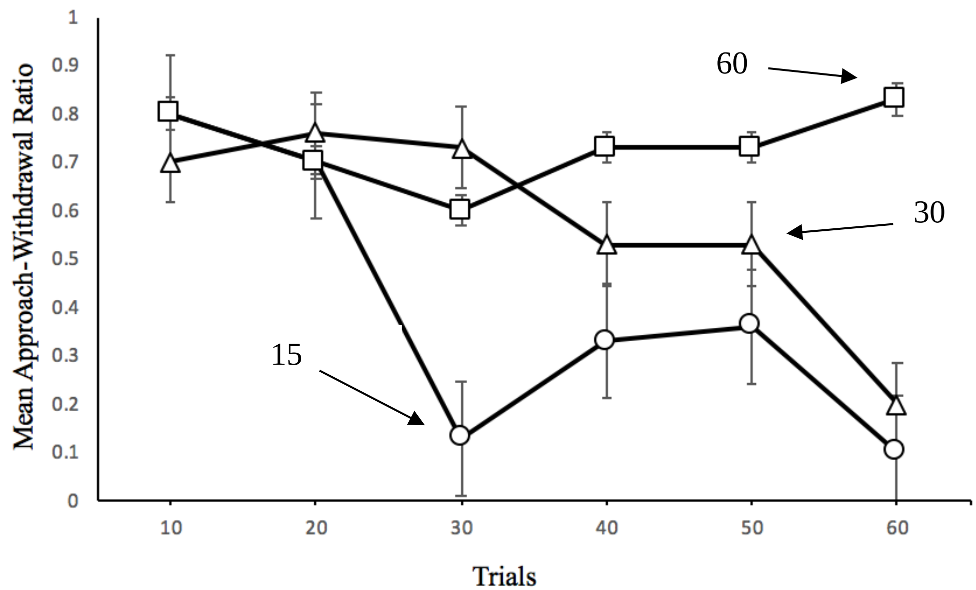

Figure 7. Mean approach-withdrawal ratios to the CS across 10-trial blocks for subjects during pairing in 15-s ITI (closed squares), 30-s ITI (open squares), and 60-s ITI conditions (closed circles). Error bars represent standard errors.